Приложение 1

Байесовские решающие правила

Для количественной оценки последствий того или иного решения вводится матрица потерь, которая сопоставляет каждой комбинации события (сигнала) s_i и решения d_j плату $\Pi_{ij} > 0$ [2, 8].

Средние потери при использовании определенного решающего правила $\delta(y)$ (способа разбиения пространства наблюдений) при усреднении по *решениям для заданного события (сигнала)* s_i равны $r_i = \sum_{j=0}^{M-1} \Pi_{ij} P(d_j \mid s_i)$, $i = 0, 1, 2, \ldots, M-1$. Эти величины называют *условными рисками*. Каждому решающему правилу $\delta(y)$, таким образом, можно сопоставить вектор потерь $r(\delta) = (r_0, r_1, \ldots, r_{M-1})$, который полностью характеризует это правило. Однако сравнение различных алгоритмов по векторным показателям качества затруднительно. Как правило, необходим переход к единому показателю качества.

Байесовские критерии предполагают случайность событий $s_0, s_1, \ldots, s_{M-1}$, для которых вводятся *априорные* вероятности $p_0, p_1, \ldots, p_{M-1}$. При известных значениях p_i достаточной характеристикой качества различения с помощью правила $\delta(y)$ служит *средний риск* (математическое ожидание потерь) $R(\delta) = \sum_{i=0}^{M-1} p_i r_i$.

Оптимальное правило δ_B называется *байесовским* и обеспечивает минимум среднего риска (байесовский риск) $R_B = \min_{\delta} R(\delta)$. Байесовский обнаружитель сравнивает отношение правдоподобия с *байесовским порогом*, и обнаруживает сигнал, если $\Lambda(y) \geq \Lambda_B$, где $\Lambda_B = (\Pi_{01} - \Pi_{00}) p_0 / (\Pi_{10} - \Pi_{11}) p_1$. Значение байесовского порога совпадает с производной $\mathrm{d}D/\mathrm{d}F$ в рабочей точке на *собственной* рабочей характеристике [1].

Критерий *идеального наблюдателя* (Зигерта) [24] минимизирует *полную вероятность ошибки* $p_e = \sum_{i=0}^{M-1} p_i \sum_{j \neq i} P(d_j \mid s_i)$. Он представляет частный случай байесовского критерия для единичных стоимостей ошибок и нулевых стоимостей правильных решений: $\Pi_{ij} = 1$ при $i \neq j$, и $\Pi_{ii} = 0$. Оптимальное различитель выбирает *наибольшую из величин* $p_i f_i(y)$, $i = 0, 1, 2, \ldots, M-1$. Указанные величины входят в *апостериорные* вероятности

$P(s_i \mid y) = p_i f_i(y) / \sum_{j=0}^{M-1} p_j f_j(y)$, поэтому различитель принимает решение по *максимуму апостериорной вероятности.*

Для задачи обнаружения сигнала полная вероятность ошибки равна $p_\Sigma = p_0 F + p_1(1-D)$, а порог Зигерта для отношения правдоподобия есть $\Lambda_S = p_0 / p_1$.

Критерий *максимального правдоподобия* предписывает выбор для заданной выборки y наибольшей из функций правдоподобия $f_i(y)$. Фактически в дополнение к критерию Зигерта полагается равенство априорных вероятностей p_i. Оптимальный различитель минимизирует *суммарную вероятность ошибки* $p_\Sigma = \sum_{i=0}^{M-1} \sum_{j \neq i} P(d_j \mid s_i) / M$. Порог Λ_{MP} для отношения правдоподобия в задаче обнаружения равен единице, а суммарная вероятность ошибки равна $p_\Sigma = (F + 1 - D)/2$.

Важным для решения задач является понятие *апостериорного риска* для данного j-го решения $R_j(\delta_j \mid y) = \sum_{i=0}^{M-1} \Pi_{ij} P(s_i \mid y)$. Он представляет средние потери по апостериорному распределению состояний (сигналов) после того, как получено наблюдение y. Апостериорные риски, как и апостериорные вероятности, являются *случайными* величинами. Вводя вектор-столбец апостериорных рисков

$$R(\delta \mid y) = (R_0(\delta_0 \mid y), R_1(\delta_1 \mid y), \ldots, R_{M-1}(\delta_{M-1} \mid y))^T,$$

можно выразить [2] средний риск в виде

$$R(\delta) = \int_Y w(y) \delta^T(y) R(\delta \mid y) \mathrm{d}y,$$

где $\delta^T(y) = (\delta_0(y), \delta_1(y), \ldots, \delta_{M-1}(y))$ – решающая вектор-функция, T – знак транспонирования, а $w(y) = \sum_{i=0}^{M-1} p_i f_i(y)$ – маргинальная плотность вероятности наблюдений.

Поскольку $w(y) \geq 0$, справедлива другая формулировка байесовского правила: оптимальный различитель выбирает для заданной выборки y решение d_j, обеспечивающее наименьшую величину апостериорного риска $R_j(\delta_j \mid y)$. Можно показать [8], что байесовское правило является *нерандомизированным*.

Подставляя апостериорные вероятности $P(s_i \mid y) = p_i f_i(y) / w(y)$, получаем выражение для апостериорный риска $R_j(\delta_j \mid y) = (1/w(y)) \sum_i \Pi_{ij} p_i f_i(y)$. Оно показывает, что ре-

шение выбирается путем сравнения *линейных комбинаций* (взвешенных сумм функций правдоподобия $f_i(y)$), которые и являются определяющими элементами задачи. Значения матрицы потерь и априорных вероятностей определяют весовые коэффициенты этих комбинаций.

Для практически важных применений значения матрицы потерь и априорные вероятности должны относительно слабо влиять на структуру оптимального алгоритма и выбор оптимального решения, и могут быть взяты произвольно [8], поскольку именно наблюдения должны быть определяющими с точки зрения принятия правильного решения. В противном случае оптимальное решение конструируется фактически на основе априорных данных о потерях и вероятностях событий.

При не очень сильных ограничениях структура байесовского правила определяет *полный класс* алгоритмов обнаружения и различения, который содержит вообще все наилучшие алгоритмы, в том числе и в условиях априорной неопределенности [8].