

Лекция №3. Байесовская теория принятия решения применительно к задаче распознавания образов

В рамках статистического подхода предполагается, что:

- образы объектов в пространстве используемых признаков представляются как реализации случайного вектора $\mathbf{x} = (x_1, \dots, x_n)^T$;
- каждое значение этого вектора $x = (x_1, \dots, x_n)^T \in R^n$ представляет образ конкретного объекта;
- любой образ может принадлежать одному из M классов объектов, в совокупности составляющих конечное множество альтернативных статистических гипотез $\Omega = \{\omega_i, i = \overline{1, M}\}$;
- количество классов и их исходные наименования считаются известными.

Для описания классов задаются **априорные вероятности** появления объектов различных классов – априорные вероятности гипотез $\omega_i : \Pr(\omega = \omega_i) = p(\omega_i), i = \overline{1, M}$ и статистическое описание вектора признаков каждого класса \mathbf{x} в виде условных плотностей распределения вероятностей: $f_{\mathbf{x}}(x/\omega_i) = p(x/\omega_i), i = \overline{1, M}$, которые называются **функциями правдоподобия классов**.

В соответствие с формулой Байеса такое статистическое описание классов эквивалентно заданию апостериорных вероятностей классов относительно полученных наблюдений x

$$p(\omega_i / x) = \frac{p(x / \omega_i) p(\omega_i)}{p(x)}, \quad i = \overline{1, M} \quad p(x) = \sum_{i=1}^M p(x / \omega_i) p(\omega_i) = \sum_{i=1}^M p(x, \omega_i)$$

Решаемая задача состоит в том, чтобы для каждого образа x выполнить действие – решение $\alpha(x)$, относящие его к тому или иному классу.

Для любого x подобные решения имеют конечное множество возможных значений $\alpha(x) \in A$, $A = \{\alpha_i, i = \overline{1, M}\}$. Множеству A соответствует разбиение пространства признаков на непересекающиеся области $\Gamma_i, i = \overline{1, M}$, $\Gamma_i \cap \Gamma_j = \emptyset, i \neq j, \bigcup_{i=1}^M \Gamma_i = R^n$, определяемые уравнениями $\alpha(x) = \alpha_i, i = \overline{1, M}$. Области $\Gamma_i, i = \overline{1, M}$ будем называть областями решений.

С учетом введенных обозначений каждое фактическое действие α_i , как значение функции $\alpha(x)$, определяет, что при обработке информации тем или иным образом установлено, что $x \in \Gamma_i$. Обозначим это следующим образом:

$$\alpha_i = \alpha(x): x \in \Gamma_i \rightarrow \omega_i.$$

Синтез решающих правил на основе различных критериев

Рассмотрим методы, позволяющие оптимизировать процесс принятия подобных решений по определенным критериям, т.е. выполнить синтез алгоритма распознавания.

Критерий минимума условного риска. Выполнение любых действий практически всегда сопряжено с ошибками. В данном случае эти ошибки возникают в случае, когда при истинности любой гипотезы ω_j будет принято решение α_i $i \neq j$. Это ведет к определенным потерям, которые выражаются количественным образом. Требуется синтезировать классификатор – алгоритм принятия решений (решающее правило), которые минимизирует ожидаемые потери от выполнения своих действий. Согласно принятой терминологии ожидаемые потери называются риском R .

В соответствии с байесовской теорией ожидаемые потери определяются на основе задания штрафных функций

$$\lambda_{ij} = \lambda(\alpha_i / \omega_j), i = \overline{1, M}, j = \overline{1, M},$$

т.е. платы за принятие решения α_i при истинности гипотез ω_j .

Вводится **функция условного риска (ФУР)**

$$r(\alpha_i / x) = \sum_{j=1}^M \lambda(\alpha_i / \omega_j) p(\omega_j / x), i = \overline{1, M}.$$

ФУР определяют суммарные потери за принятие решения α_i при получении данных x . Тогда **общий ожидаемый риск** можно представить как математическое ожидание условного риска относительно распределения всех возможных значений СВ x

$$R = \int r(\alpha(x)/x) p(x) dx = \sum_{i=1}^M \int_{\Gamma_i} r(\alpha_i/x) p(x) dx = \sum_{i=1}^M \sum_{j=1}^M \int \lambda(\alpha_i/\omega_j) p(\omega_j, x) dx. \quad (1)$$

Поскольку общий риск определяется как интеграл от взвешенного неотрицательной функцией $p(x)$ условного риска, то из (1) следует, что для того, чтобы минимизировать ожидаемые потери нужно каждый раз при получении на вход классификатора конкретного значения x минимизировать условный риск.

Другими словами, если для каждого x решение $\alpha(x)$ выбрано таким образом, что величина $r(\alpha(x)/x)$ является минимальной из всех возможных, то и величина общего риска будет минимальной, т.е. для $\forall x$ должно выполняться

$$r(\alpha(x)/x) = \min\{r(\alpha_1/x), \dots, r(\alpha_M/x)\} \rightarrow \min R.$$

В итоге, можно сформулировать следующее правило принятия решения в ходе распознавания: выбор в пользу гипотезы ω_i для полученного образа x осуществляется, если

$$\omega_i : r(\alpha_i/x) \leq r(\alpha_j/x), \quad j = \overline{1, M}, \quad i \neq j. \quad (2)$$

Полученное решающее правило называется **правилом минимума условного риска** (МУР).

Рассмотрим реализацию правила МУР для ситуации двух классов. В этом случае

$$r(\alpha_1/x) = \lambda_{11}p(\omega_1/x) + \lambda_{12}p(\omega_2/x),$$

$$r(\alpha_2/x) = \lambda_{21}p(\omega_1/x) + \lambda_{22}p(\omega_2/x).$$

Тогда при поступлении образа x осуществляется выбор класса ω_1 , если $r(\alpha_1/x) \leq r(\alpha_2/x)$, или класса ω_2 , если $r(\alpha_1/x) > r(\alpha_2/x)$. Используя представленные соотношения и, объединяя эти неравенства в общее выражение, получим

$$(\lambda_{21} - \lambda_{11})p(\omega_1/x) \begin{matrix} > \\ < \end{matrix}^{\omega_1}_{\omega_2} - (\lambda_{22} - \lambda_{12})p(\omega_2/x)$$

или

$$\frac{(\lambda_{21} - \lambda_{11})p(x/\omega_1)p(\omega_1)}{p(x)} \begin{matrix} > \\ < \end{matrix}^{\omega_1}_{\omega_2} \frac{(\lambda_{12} - \lambda_{22})p(x/\omega_2)p(\omega_2)}{p(x)}.$$

Обычно потери в случае ошибки больше, чем при правильном распознавании и $\lambda_{21} - \lambda_{11} > 0, \lambda_{12} - \lambda_{22} > 0$. Тогда, сокращая общий сомножитель, неравенства окончательно удобно представить в виде

$$l(x) = \frac{p(x/\omega_1)}{p(x/\omega_2)} \begin{matrix} > \\ < \end{matrix}^{\omega_1}_{\omega_2} \frac{(\lambda_{12} - \lambda_{22})p(\omega_2)}{(\lambda_{21} - \lambda_{11})p(\omega_1)} = l_0. \quad (3)$$

Величина $l(x)$ в (3) называется **отношением правдоподобия**. Алгоритм принятия решения сводится к сравнению отношения правдоподобия $l(x)$ с порогом l_0 , который зависит от штрафных функций и априорных вероятностей гипотез, определяющих априорные сведения о появлении классов образов. Области решений в пространстве признаков определяются в этом случае как

$$\alpha_1 : x \in \Gamma_1 : l(x) \geq l_0, \quad \alpha_2 : x \in \Gamma_2 : l(x) < l_0,$$

а разделяющая граница между ними – уравнением $l(x) = l_0$.

Критерий максимума апостериорной вероятности. Важным частным случаем являются использование симметричных штрафных функций с нулевой платой за правильное решение

$$\lambda_{ij} = \lambda(\alpha_i / \omega_j) = \begin{cases} 0, & i = j \\ 1, & i \neq j \end{cases} \quad i = \overline{1, M}, \quad j = \overline{1, M}.$$

Использование штрафных функция такого вида фактически означает, что отсутствуют обоснованные рекомендации относительно того, как разные ошибочные решения могут повлиять на ожидаемые потери. Поэтому влияние всех ошибочных решений признается одинаковым, а плата за правильное решение приравнивается нулю.

В этом случае выражение для условного риска преобразуется следующим образом:

$$r(\alpha_i / x) = \sum_{\substack{j=1 \\ j \neq i}}^M p(\omega_j / x) = 1 - p(\omega_i / x), \quad i = \overline{1, M}.$$

Отсюда следует, что принятие решения в соответствии с критерием минимума условного риска в данном случае эквивалентно принятию решения в соответствии с максимумом апостериорной вероятности гипотезы о появлении объекта одного из классов. Байесовское решающее правило в случае многих классов будет иметь вид системы неравенств для апостериорных вероятностей гипотез

$$\omega_i : p(\omega_i / x) \geq p(\omega_j / x), \quad j = \overline{1, M}, \quad i \neq j \quad (4a)$$

или системы неравенств для отношений правдоподобия

$$\omega_i : l_{ij}(x) = \frac{p(x/\omega_i)}{p(x/\omega_j)} \underset{\omega_2}{\overset{\omega_1}{>}} l_{ij} = \frac{p(\omega_j)}{p(\omega_i)}, \quad j = \overline{1, M}, \quad i \neq j. \quad (4б)$$

При проверке гипотез в случае двух классов алгоритм преобразуется следующим образом:

$$l(x) = \frac{p(x/\omega_1)}{p(x/\omega_2)} \underset{\omega_2}{\overset{\omega_1}{>}} \frac{p(\omega_2)}{p(\omega_1)} = l_0. \quad (5)$$

Полученные решающие правила будем называть правилами **максимума апостериорной вероятности** (МAB). Их особенностью фактически является отказ от обоснования штрафных функций и использование при вычислении порога принятия решения только априорных вероятностей гипотез, т.е. требуется меньше исходной информации по сравнению с МУР.

Критерий максимального правдоподобия. Еще один важный частный случай возникает в ситуации, когда, в дополнение к отсутствию обоснованных штрафных функций, априорные вероятности классов одинаковы $p(\omega_i) = 1/M$, $i = \overline{1, M}$ или их значения просто неизвестны. В этом случае полученные решающие правила (4) естественным образом преобразуются к виду

$$\omega_i: p(x/\omega_i) \geq p(x/\omega_j), \quad j = \overline{1, M}, \quad i \neq j, \quad (6a)$$

или к виду

$$\omega_i: l_{ij}(x) = \frac{p(x/\omega_i)}{p(x/\omega_j)} \underset{\omega_2}{\overset{\omega_1}{>}} l_{ij} \equiv 1, \quad j = \overline{1, M}, \quad i \neq j. \quad (6b)$$

При распознавании двух классов решающее правило выглядит следующим образом:

$$l(x) = \frac{p(x/\omega_1)}{p(x/\omega_2)} \underset{\omega_2}{\overset{\omega_1}{>}} l_0 \equiv 1.$$

Данные алгоритмы реализуют критерий максимума правдоподобия (МП). Далее будем их также называть алгоритмами МП. При их использовании реализуется минимальный объем исходной информации, необходимой для синтеза алгоритма распознавания.

Обобщенная структура решающего правила. Понятие разделяющей функции.

Исходя из результатов проведенного выше синтеза алгоритмов, можно предложить обобщенную структуру байесовского классификатора, определяющую каноническую форму представления с использованием так называемых **разделяющих функций**.

В качестве разделяющих функций будем рассматривать набор функций $g_i(x), i = \overline{1, M}$, на основе которых может быть определено соответствие вектора признаков одному из классов по правилу

$$\omega_i : g_i(x) \geq g_j(x), \quad j = \overline{1, M}, \quad i \neq j,$$

т.е. отнесение образа объекта к тому или иному классу осуществляются на основе выбора максимума среди соответствующих значений разделяющих функций.

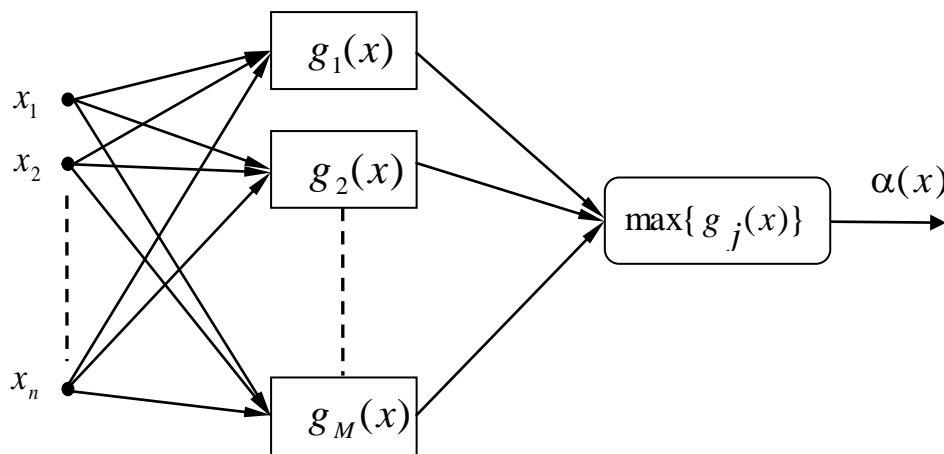


Рис.1. Обобщенная структура решающего правила

В схеме рис.1 отображается общая сущность ранее рассмотренных решающих правил, используемых в процессе распознавания. Решение в любом из них в итоге формируется как индекс класса, у которого значение разделяющей функции максимально. Для рассмотренных ранее трех вариантов алгоритмов разделяющие функции можно определить следующим образом:

$$g_i(x) = \begin{cases} -r(\alpha_i / x), \\ p(\omega_i / x); p(\omega_i)p(x / \omega_i), & i = \overline{1, M}. \\ p(x / \omega_i), \end{cases}$$

Очевидно, что, если $g_i(x), i = \overline{1, M}$ являются разделяющими функциями, то **любая монотонно возрастающая функция** $g'(x) = f(g(x))$ от них позволяет получить эквивалентное решающее правило для системы новых функций $g'_i(x), i = \overline{1, M}$.

Особенно удобно в таких случаях использовать преобразование на основе натурального логарифма $g'_i(x) = \ln(g_i(x)), i = \overline{1, M}$, поскольку при этом произведения и степени преобразуются в суммы, что может существенно упростить математический анализ и процесс вычислений.

В частном случае распознавания двух классов с использованием разделяющих функций реализуется следующее обобщенное решающее правило:

$$g_1(x) \underset{\omega_2}{\overset{\omega_1}{>}} g_2(x), \text{ что эквивалентно } g(x) = g_1(x) - g_2(x) \underset{\omega_2}{\overset{\omega_1}{>}} 0.$$

Это означает, что классификатор двух классов удобно представить как алгоритм, вычисляющий единственную разделяющую функции. Отметим, что в случае двух классов уравнение $g_1(x) - g_2(x) = 0$ определяет границу между областями решений Γ_1, Γ_2 .

При использовании, например, критерия МАВ, разделяющая функция может иметь вид

$$g(x) = p(\omega_1 / x) - p(\omega_2 / x)$$

или в эквивалентном виде

$$g'(x) = \ln p(\omega_1 / x) - \ln p(\omega_2 / x) = \ln \frac{p(x / \omega_1)}{p(x / \omega_2)} + \ln \frac{p(\omega_1)}{p(\omega_2)}.$$

Следует отметить, что при выполнении логарифмического преобразования для общей формы классификатора двух гипотез

$$l(x) = \frac{p(x / \omega_1)}{p(x / \omega_2)} \underset{\omega_2}{\overset{\omega_1}{>}} = l_0 ,$$

где l_0 – порог, задаваемый для каждого из алгоритмов МУР, МАВ, МП, логарифмирование выполняется и для порога. Т.е. общий вид алгоритма классификации будет иметь вид

$$g'(x) = \ln l(x) - l'_0 = g''(x) - l'_0 \underset{\omega_2}{\overset{\omega_1}{>}} 0. \quad g''(x) = \ln l(x), \quad l'_0 = \ln l_0.$$

Анализ решающих правил. Расчет вероятностей ошибок распознавания

Важнейшим этапом разработки алгоритмов распознавания, как и любых алгоритмов обработки информации, является этап анализа ранее синтезированного алгоритма. Несмотря на то, что алгоритмы, синтезированные на основе введенных критериев, являются оптимальными и улучшить их работу в рамках исходной модели нельзя, их анализ с точки зрения ожидаемых потерь и ошибок позволяет разработчику сформировать предметные практические рекомендации. Эти рекомендации могут касаться возможности использования полученного алгоритма в принципе, необходимости расширения перечня используемых признаков, повышения точности их измерения, использования новых физических датчиков и т.п.

Кроме того, в ходе анализа синтезированных алгоритмов важной задачей являются анализ вида разделяющих функций (линейные, нелинейные) и формы областей принятия решений $\Gamma^{(i)}$, $i = \overline{1, M}$, что зачастую позволяет сформировать конструктивные рекомендации относительно возможных упрощений алгоритмов.

Базовыми показателями являются **вероятности правильного распознавания** объектов данного класса $P_c^{(i)}$, $i = \overline{1, M}$ и **вероятности ошибки** при распознавании объектов данного класса $\varepsilon_{er}^{(i)}$, $i = \overline{1, M}$.

Указанные показатели определяются соотношениями

$$P_c^{(i)} = 1 - \varepsilon_{er}^{(i)} = \int_{\Gamma_i} p(x/\omega_i) dx, \quad \varepsilon_{er}^{(i)} = \int_{\bar{\Gamma}_i} p(x/\omega_i) dx, \quad i = \overline{1, M}. \quad (7)$$

На их основе рассчитывается общая суммарная ошибка распознавания, усредненная с учетом априорных вероятностей классов

$$E_s = \sum_{i=1}^M p(\omega_i) \varepsilon_{er}^{(i)} = 1 - \sum_{i=1}^M p(\omega_i) P_c^{(i)}.$$

Помимо этого, используется такая характеристика как **матрица ошибок распознавания** классов, в которой недиагональные элементы составляют вероятности перепутывания классов, а на главной диагонали находятся вероятности правильного распознавания классов,

$$M_{er} = \|\varepsilon_{ij}\|, \quad \varepsilon_{ij} = \int_{\Gamma_j} p(x/\omega_i) dx, \quad \sum_{\substack{j=1 \\ j \neq i}}^M \varepsilon_{ij} = \varepsilon_{er}^{(i)}, \quad \varepsilon_{ii} = P_c^{(i)} = 1 - \sum_{\substack{j=1 \\ j \neq i}}^M \varepsilon_{ij}.$$

При распознавании двух классов вероятности ошибок разделяют на ошибку первого рода и ошибку второго рода. Они определяются как

$$\alpha = \int_{\Gamma_2} p(x/\omega_1) dx, \quad \beta = \int_{\Gamma_1} p(x/\omega_2) dx. \quad (8)$$

В приведенных ниже рисунках даны примеры двух вариантов разделяющих функций для МАВ в виде $p(\omega_i)p(x/\omega_i)$ и в виде функций апостериорных вероятностей гипотез $p(\omega_i/x)$ относительно x , а также областей принятия решений, выделенных вертикальными линиями. Рассматривается случай распознавания трех классов образов с различными априорными вероятностями. Кроме того, проводится визуализация ошибок перепутывания двух классов ε_{ij} , которые определяются площадью закрашенных областей под кривыми функций правдоподобия классов.

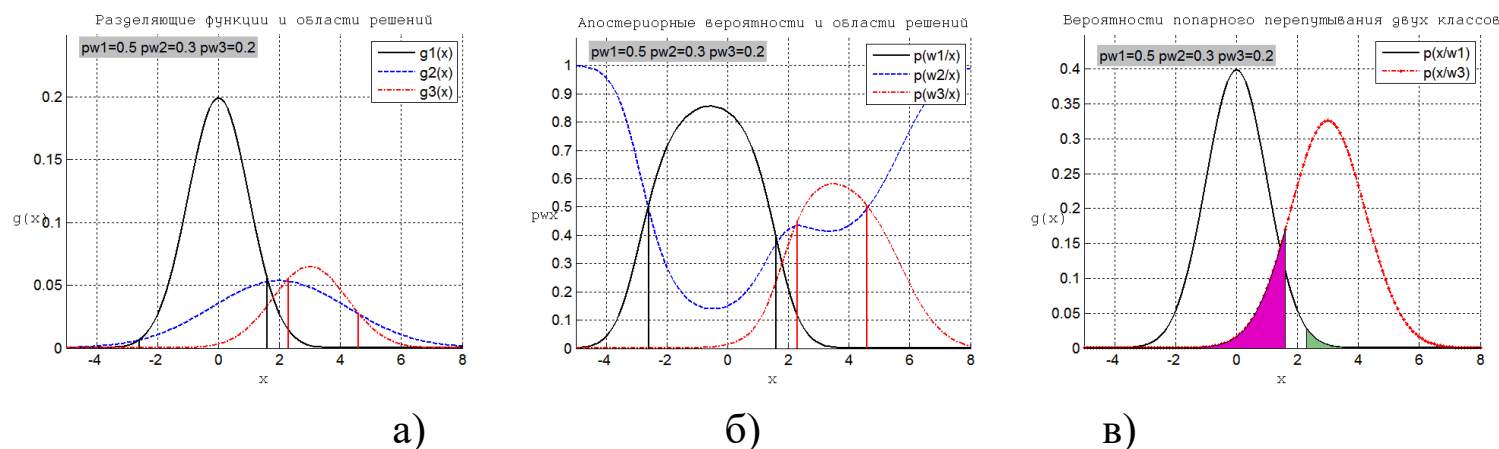


Рис.2. Разделяющие функции (а), апостериорные вероятности (б), а также вероятности ошибок перепутывания первого и третьего классов (в)

Мы видим, что различные формы задания разделяющей функции в данном случае дают одинаковый результат с точки зрения локализации областей принятия решений, выделенных вертикальными линиями.

Для вероятностей перепутывания первого и третьего классов (рис.2,в) промежуточная не закрашенная область соответствует центральной подобласти решения второго класса, которая в целом состоит из трех подобластей, размещенных слева, по центру и справа на оси ОХ.

Для случая двух классов пример получения аналогичных графиков, представлен на рис.3. Закрашенные области здесь отображают вероятности ошибок первого и второго рода.

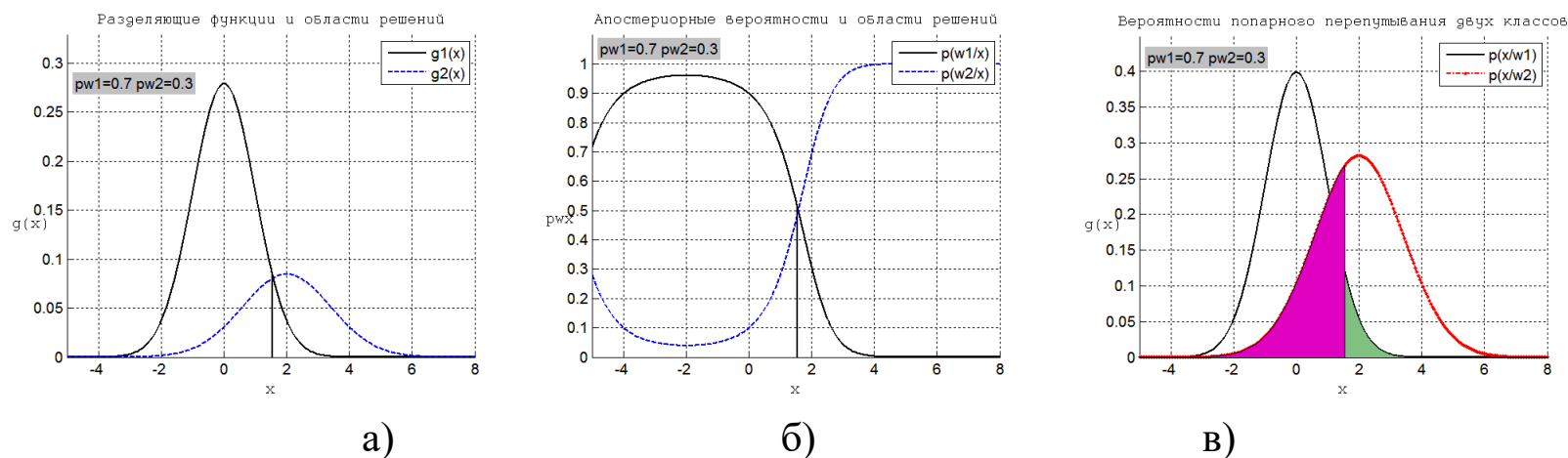


Рис.3. Разделяющие функции (а), апостериорные вероятности (б), а также вероятности ошибок первого и второго рода (в)

При анализе эффективности алгоритмов распознавания на основе минимума условного риска вероятности ошибочных решений при наличии штрафных функций легко пересчитываются в ожидаемые потери. Для этого преобразуем выражение для общего риска следующим образом:

$$R = \sum_{i=1}^M \sum_{j=1}^M \lambda(\alpha_i / \omega_j) p(\omega_j) \int_{\Gamma_i} p(x / \omega_j) dx = \sum_{i=1}^M \sum_{j=1}^M \lambda(\alpha_i / \omega_j) \varepsilon_{ji} p(\omega_j).$$

Подставим в выражение для общего риска симметричную функцию потерь. В этом случае общий риск преобразуется к виду

$$R = \sum_{i=1}^M \int_{\Gamma_i} [1 - p(\omega_i / x)] p(x) dx = 1 - \sum_{i=1}^M p(\omega_i) \int_{\Gamma_i} p(x / \omega_i) dx = 1 - \sum_{i=1}^M p(\omega_i) P_c^{(i)} = E_s.$$

Каждый интеграл под знаком суммы определяет вероятность того, что $x \in \Gamma_i$, если анализируемый объект действительно принадлежит классу ω_i , т.е. вероятность правильного распознавания объектов этого класса $P_c^{(i)}$, $i = \overline{1, M}$. Соответственно всё выражение дает усредненную по априорным вероятностям классов суммарную вероятность ошибки распознавания E_s .

Таким образом, применение критерия МАВ обеспечивает минимизацию суммарной ошибки. Поэтому в литературе данный критерий еще называют **критерием идеального наблюдателя**.

Конкретное определение представленных показателей может производиться различными способами.

Первый способ предполагает проведение прямого интегрирования исходных плотностей $p(x/\omega_j)$, $i=\overline{1,M}$ на основе (7) или (8) по областям решений $\Gamma^{(i)}$, $i=\overline{1,M}$. Однако такое интегрирование для случая, когда размерность признакового пространства больше 1, весьма затруднительно.

Второй способ предполагает проведение анализа одномерных распределений разделяющих функций с учетом вида этих функций $p(g_j/\omega_j)$, $i=\overline{1,M}$, а также использования приближений для построения этих распределений. Обычно после логарифмирования разделяющие функции представляются в виде сумм определенного числа СВ, что часто дает основания применить для них центральную предельную теорему и использовать приближенно гауссовское распределение для $p(g''/\omega_i)$, $g''(x) = \ln l(x)$. Это весьма упрощает задачу, как далее мы увидим.

Полученные распределения используются для расчета ошибок. Для случая двух классов при использовании слагаемого разделяющей функции $g''(x) = \ln l(x)$, являющейся скалярной величиной логарифма отношения правдоподобия (ЛОП), вероятности ошибок определяются как одномерные интегралы вида

$$\alpha = \int_{-\infty}^{l_0} p(g''/\omega_1) dg'', \quad \beta = \int_{l_0}^{\infty} p(g''/\omega_2) dg''.$$

Третий способ аналитического расчета ошибок базируется на использовании **принципа гарантированного результата**, в основе которого лежит использование различного рода верхних границ для вероятностей ошибок.

К ним относится граница Чернова, которая может быть рассчитана при распознавании двух классов как для каждого класса в отдельности, так и для суммарной ошибки в виде

$$E_s = p(\omega_1)\alpha + p(\omega_2)\beta \leq p(\omega_1)^{1-t} p(\omega_2)^t \exp(-\mu(t)), \quad \mu(t) = -\ln \int p(x/\omega_1)^{1-t} p(x/\omega_2)^t dx.$$

где t — скалярная величина, лежащая в интервале от нуля до единицы. Оптимальное значение t удовлетворяет уравнению

$$-\frac{d\mu(t)}{dt} = \ln p(\omega_1)/p(\omega_2).$$

Величина $\mu(1/2)$ называется расстоянием Бхатачария. Его также используют для оценки степени разделимости двух классов образов.

Для случая многих классов с использованием границы Чернова может быть оценена верхняя граница E_s на основе следующего неравенства:

$$E_s < \sum_{j=1}^M \sum_{i>j}^M E_{s,ij},$$

где $E_{s,ij}$ — верхние границы, оцениваемые для случая попарного распознавания классов.

Четвертый и, пожалуй, самым простой способ предполагает проведение оценки вероятностей ошибок путем прямого статистического имитационного моделирования алгоритма распознавания, в ходе которого проводится подсчет количества возникающих ошибок и усреднение результатов по числу K выполненных реализаций моделирования. Подобный подход не требует использования сложных интегральных преобразований, но требует внимательного отношения к анализу уровня достоверности получаемых результатов статистического моделирования, особенно, в случаях, когда оцениваются весьма малые вероятности ошибок или близкие к единице вероятности правильного распознавания классов.

Поэтому при оценке ошибок в ходе статистического моделирования используют метод доверительных интервалов. Требуется найти число испытаний, при котором

$$P\left[\left|P_c^{(i)} - \tilde{P}_c^{(i)}\right| < d_g\right] = 1 - \gamma,$$

где d_g — доверительный интервал; γ — уровень значимости. Конечное уравнение, связывающее d_g и K , имеет вид

$$d_g = t_{kp}(\gamma) \sqrt{P_c^{(i)}(1 - P_c^{(i)})/K}, \quad K = \frac{t_{kp}^2(\gamma) P_c^{(i)}(1 - P_c^{(i)})}{d_g^2}, \quad 2\Phi(t_{kp}(\gamma)) - 1 = 1 - \gamma,$$

$$t_{kp}(\gamma) = \Phi^{-1}((2 - \gamma)/2), \quad \Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{x^2}{2}} dx$$

Для того, чтобы избавиться от неизвестного значения $P_c^{(i)}$, можно заменить произведение $P_c^{(i)}(1 - P_c^{(i)})$ на величину $\max_p P_c^{(i)}(1 - P_c^{(i)}) = 0,25$. Тогда окончательно получим выражение для гарантированных значений d_g и K

$$d_p^z \geq \frac{t_{kp}(\gamma)}{2\sqrt{K}}, \quad K \geq \frac{t_{kp}^2(\gamma)}{4d_g^2}.$$

В приведенной ниже таблице 1 представлены рекомендуемые объемы испытаний – значения K , полученные для уровня значимости $\gamma = 0.05$ и доверительного интервала, выбираемого исходя из 25% и 10% - го значения доверительного интервала относительно оцениваемого уровня вероятности. Величина $t_{kp}(\gamma = 0.05) = 1.96$.

Таблица 1. Рекомендуемые объемы статистических испытаний

Значения оцениваемой вероятности ошибки $(1 - P_c^{(i)})$	Количество испытаний K	
	$d_g = 0.25(1 - P_c^{(i)})$	$d_g = 0.1(1 - P_c^{(i)})$
0.1	5.532e+02	3.457e+03
0.01	6.085e+03	3.803e+04
0.001	6.140e+04	3.837e+05

Представленные данные свидетельствуют о том, что для оценки малых вероятностей ошибки или, что, эквивалентно, близких к единице вероятностей правильного распознавания, требуется использовать количество испытаний почти **на два порядка** превышающее величину, обратную оцениваемой малой вероятности ошибки.