Лекция 8

1. Общесистемные вопросы теории управления

Термины геометрический, метрический, детерминистский подход к решению задач распознавания образов используются во многих работах для обозначения разнообразных методов и алгоритмов, основанных на отказе от статистических моделей данных и постулировании некоторых гипотез о структуре классов и геометрических свойствах образов, представляющих эти классы в многомерном пространстве признаков. К таким гипотезам, как уже упоминалось, относится, прежде всего, гипотеза компактности, согласно которой схожие в метрическом смысле объекты гораздо чаще лежат в одном классе, чем в разных; при этом классы образуют компактно локализованные подмножества в пространстве объектов, а границы между ними имеют достаточно простую форму. Другой важной особенностью детерминистского подхода является наличие набора прецедентов в виде обучающей выборки, на основе которой строится решающее правило, имеющее заранее заданную структуру.

Важно также обратить внимание, что во многих случаях задача может быть решена в рамках другого подхода, например, статистического. При этом могут быть получены близкие, а иногда и совпадающие результаты. Такие аналогии будут нами далее комментироваться. Следует также отметить, что в приведенных ниже иллюстративных примерах при необходимости обеспечить задание контролируемых наборов исходных данных, мы будем использовать статистические алгоритмы генерации, которые описаны в предыдущих главах. Но интерпретация получаемых на этой основе результатов будет отличаться. Например, вместо термина оценка вероятности ошибок распознавания мы будем использовать термин частость ошибок (отношение числа ошибок к общему числу испытаний).

Лекция 8.

Распознавание образов с использованием функций расстояния

Распознавание образов с использование функций расстояния (мер близости) является одной из первых идей, реализованных в автоматических распознающих системах. Этот относительно простой метод основан на предположении о том, что образы в пределах каждого класса имеют разумную степень изменчивости, при этом между собой классы образов пересекаются слабо.

Решаемая задача, как и ранее в статистической теории, 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 решений Γ_i, $i = \overline{1,M}$, $\Gamma_i \bigcap \Gamma_j = \varnothing$, $i \neq j$, $\bigcup_{i=1}^{M} \Gamma_i = R^n$, определяемые уравнениями $\alpha(x) = \alpha_i$, $i = \overline{1,M}$. Каждое фактическое действие α_i, как значение функции $\alpha(x)$, означает, что при обработке информации установлено, что $x \in \Gamma_i$, т.е. $\alpha_i = \alpha(x):\ x \in \Gamma_i \to \omega_i$.

Вещественная неотрицательная функция $d(x^{(i)}, x^{(j)})$ образов $x^{(i)} \in R^n$, $x^{(j)} \in R^n$ называется метрикой, если для $\forall x^{(i)}, x^{(j)}$ выполняется

1) $d(x^{(i)}, x^{(i)}) = 0$;

2) $d(x^{(i)}, x^{(j)}) = d(x^{(j)}, x^{(i)})$;

3) $d(x^{(i)}, x^{(j)}) = 0 \to x^{(i)} \equiv x^{(j)}$;

4) $d(x^{(i)}, x^{(j)}) \leq d(x^{(i)}, x^{(k)}) + d(x^{(k)}, x^{(j)})$;

Иногда два последних свойства игнорируются. В этом случае, функция $d(x^{(i)}, x^{(j)})$ называется функцией расстояния. Она не является метрикой в строгом смысле.

В качестве наиболее часто употребляемых функций расстояния (ФР) можно указать Евклидово расстояние, Махаланобисово расстояние, l_p - норму (метрику Минковкого):

$$d_E(x^{(i)}, x^{(j)}) = \left[\sum_{s=1}^{n} (x_s^{(i)} - x_s^{(j)})^2\right]^{\frac{1}{2}} = \left[(x^{(i)} - x^{(j)})^T (x^{(i)} - x^{(j)})\right]^{\frac{1}{2}},$$

$$d_M(x^{(i)}, x^{(j)}) = \left[\sum_{s=1}^{n}\sum_{t=1}^{n} \bar{c}_{ts} (x_s^{(i)} - x_s^{(j)})(x_t^{(i)} - x_t^{(j)})\right] = \left[(x^{(i)} - x^{(j)})^T W^{-1} (x^{(i)} - x^{(j)})\right]^{\frac{1}{2}},$$

$$d_{lp}(x^{(i)}, x^{(j)}) = \left[\sum_{s=1}^{n} \left|x_s^{(i)} - x_s^{(j)}\right|^p\right]^{\frac{1}{p}},$$

а также ряд других.

Для построения алгоритмов, основанных на прямом использовании метрик или функция расстояния (метрических алгоритмов) в каждом классе задается один или несколько эталонных образов. Процесс получения эталонных образов будет нами рассмотрен несколько позже.

1. Решающее правило при наличии одного эталонного описания в каждом классе

Рассмотрим M классов образов, каждый из которых представлен своим одним эталонным образом ω_i: $z^{(i)}$, $i=\overline{1,M}$. Такая постановка является простейшей, но она часто встречается на практике, если известна исходная, не искаженная (эталонная) версия образа каждого класса, например, в задаче машинного распознавания печатных символов и логотипов.

Классификатор, построенный по принципу минимума функции расстояния, в данном случае определяет принадлежность x по следующему правилу:

$$\omega_i : d_E^{(i)}(x,z^{(i)}) \le d_E^{(j)}(x,z^{(j)}),\quad j=\overline{1,M},\quad i\ne j\,.$$

Для определенности, рассмотрим решение задачи при использовании Евклидова расстояния. Тогда расстояния между любым новым образом x и эталонным образом каждого класса определяется как

$$d_E^{(i)}(x,z^{(i)}) = \left[(x-z^{(i)})^T(x-z^{(i)})\right]^{\frac{1}{2}},\quad i=\overline{1,M}\,.$$

Используем вместо функций $d_E^{(i)}(x,z^{(i)})$ их квадраты, тогда получим

$$D_E^{(i)} = \left[d_E^{(i)}(x,z^{(i)})\right]^2 = (x-z^{(i)})^T(x-z^{(i)}) = x^Tx - 2x^Tz^{(i)} + z^{(i),T}z^{(i)} =$$

$$= x^Tx - 2\left(x^Tz^{(i)} - \frac{1}{2}z^{(i),T}z^{(i)}\right),\ i=\overline{1,M}\,.$$

Анализ последнего выражения показывает, что выбор минимального расстояния эквивалентен выбору максимального значения функций

$$\omega_i : g_E^{(i)}(x) \ge g_E^{(j)}(x),\quad j=\overline{1,M},\quad i\ne j\,,$$

$$g_E^{(i)}(x) = x^Tz^{(i)} - \frac{1}{2}z^{(i),T}z^{(i)},\ i=\overline{1,M}\,.$$

Таким образом, функции $g_E^{(i)}(x)$ являются разделяющими функциями линейного вида. При этом структура полученного решающего правила в целом близка структуре решающего правила, ранее полученного в задаче статистического распознавания ГСВ с одинаковыми матрицами ковариации, имеющими диагональный вид (независимые признаки).

Введем следующие обозначения $w^{(i)} = (z_1^{(i)},\dots,z_n^{(i)})$, $w_0^{(i)} = -0.5 z^{(i),T} z^{(i)}$. Тогда разделяющие функции можно представить в виде

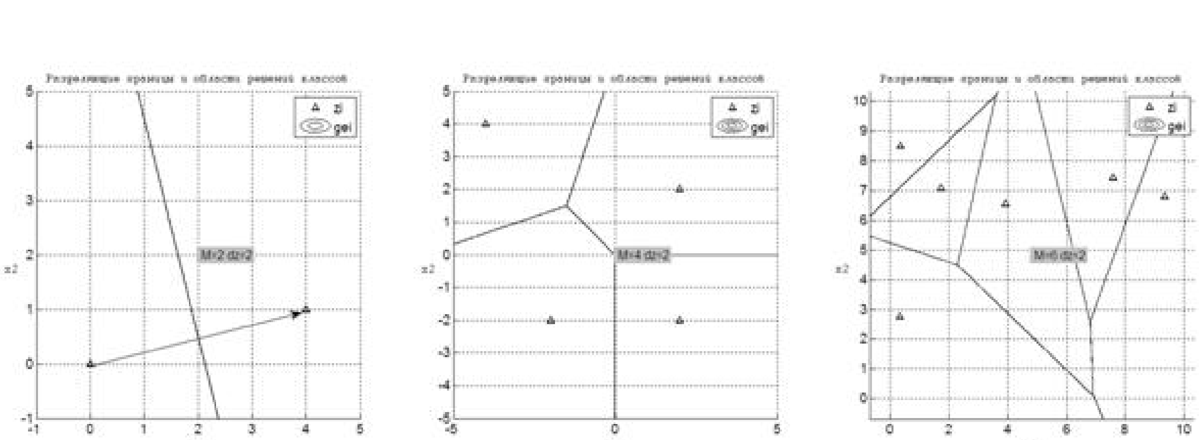
$$g_E^{(i)}(x) = x^T w^{(i)} + w_0^{(i)}, \quad i = \overline{1,M}.$$

Форма границы между двумя классами, определяется исходя из уравнения

$$g_E^{(i)}(x) - g_E^{(j)}(x) = x^T (z^{(i)} - z^{(j)}) + (w_0^{(i)} - w_0^{(j)}) = 0.$$

Это означает, что граница областей принятия решений Γ_i и Γ_j, если классов только два имеет вид гиперплоскости, которая ориентирована перпендикулярно относительно прямой, соединяющей точки $z^{(i)}, z^{(j)}$ и проходит через срединную точку этой прямой. В случае, когда классов больше двух, области решений имеют кусочно-линейные границы, образованные разделяющими функциями при попарном сравнении классов. Границы областей решений являются геометрическими местами точек, равноудаленных от прямых, соединяющих эталоны различных классов.

Для иллюстрации этих положений может быть использована программа, обеспечивающая визуализацию областей решений для двумерного признакового пространства в зависимости от расположений эталонных образов классов. На рис.1а показан пример построения с использованием представленной программы границы и соответствующих областей решений для двух классов вместе с дорисованной вручную стрелкой, отображающей вектор разности между $z^{(i)}, z^{(j)}$. На рис. 1,б,в показаны примеры построения границ четырех классов и шести классов, а также областей принятия решений, образуемых комбинациями фрагментов границ между различными парами классов.



а) б) в)

Рис.1. Примеры построения областей решений нескольких классов

2. Решающее правило при наличии нескольких эталонных описаний в каждом классе

Еще одна характерная ситуация возникает, когда описание каждого класса включает конечное и для определенности одинаковое количество эталонных образов в виде множества ω_i: $Z^{(i)} = \left\{z^{(i,k)}, k = \overline{1:L}\right\}$, $i = \overline{1,M}$, являющихся стандартными версиями одного и того же образа. В этом случае используют классификаторы следующего вида:

$$\omega_i : d_E^{(i)}\left(x, Z^{(i)}\right) \le d_E^{(j)}\left(x, Z^{(j)}\right), \quad j = \overline{1,M}, \quad i \ne j .$$

$$d_E^{(i)}\left(x, Z^{(i)}\right) = \min_k \left[d_E^{(i)}\left(x, z^{(i,k)}\right)\right], i = \overline{1,M}$$

При использовании Евклидова расстояния, следую ранее рассмотренной процедуре, можно показать, что решающее правило (6.2) эквивалентно следующему:

$$\omega_i : g_E^{(i)}(x) \ge g_E^{(j)}(x), \quad j = \overline{1,M}, \quad i \ne j ,$$

$$g_E^{(i)}(x) = \max_k \left[x^T z^{(i,k)} - \frac{1}{2} z^{(i,k),T} z^{(i,k)}\right], \; i = \overline{1,M} .$$

Анализируя полученный результат, можно увидеть, что границы областей решений в данном случае также являются геометрическими местами точек, равноудаленных от прямых, соединяющих эталоны различных классов. Полученная при этом структура классификатора эквивалента структуре классификатора при использовании одного эталонного описания из общей совокупности в отдельности с последующим объединением результатов распознавания для всех эталонов, входящих в каждые класс. Например, на рис. 1,б представлены области решений, полученные для четырех классов с единственными эталонами $z^{(1)} = (2,2)^T, z^{(2)} = (4,-4)^T, z^{(3)} = (2,-2)^T, z^{(4)} = (-2,-2)^T$. Если, теперь, имеется два класса, представленных двумя эталонными векторами каждый

$$\omega_1 : Z^{(1)} = \left\{z^{(1,1)} = (2,2)^T, z^{(1,2)} = (-2,-2)^T\right\},$$

$$\omega_2 : Z^{(2)} = \left\{z^{(2,1)} = (-4,4)^T, z^{(2,2)} = (2,-2)^T\right\},$$

(соответствующие точки показаны на рис.1,б), то области решений можно получить, объединив в диагональных направлениях области решений, ранее полученные для четырех классов с одним эталоном $\Gamma'_1=\Gamma_1\cup\Gamma_4$ $\Gamma'_2=\Gamma_2\cup\Gamma_3$.

Проводя аналогию с ранее рассмотренными статистическими алгоритмами, естественно обратить внимание на схожесть реализуемого в данном случае подхода с подходом метода k- соседей. Если не рассматривать возможности последнего для получения непараметрических оценок плотностей распределения вероятностей, то становится очевидным практическая идентичность структуры решающих правил. Так, например, использование одного эталона в каждом классе для метрического алгоритма эквивалентно использованию правила одного ближайшего соседа. Поэтому алгоритмы, реализованные методом k-соседей относят также и к метрическим алгоритмам распознавания.

3. Обучение метрических алгоритмов

Во многих случаях на практике отсутствует возможность выбора эталонов на основе исходных предположений об их виде. В подобных случаях для каждого класса образов может быть задана обучающая выборка $X^{N_i}=\{x^{(i,1)},\ldots,x^{(i,N_i)}\}$, $i=\overline{1,M}$, на основе которой можно сформировать эталоны, используемые при построении метрических алгоритмов. При этом задача состоит в том, чтобы путем прореживания обучающих данных выбрать в качестве эталонов наиболее информативные образы (опорные образы), а также, по возможности, сократить количество используемых эталонов в интересах сокращения объема вычислений на этапе принятия решений и повышения быстродействия обработки.

Обычно обучающие данные не являются равноценными. Среди них могут находиться типичные представители классов, которые и целесообразно использовать в качестве эталонов. Кроме того, могут встречаться неинформативные, или периферийные образы. Они плотно окружены другими объектами того же класса. Если их удалить из выборки, это практически не отразится на качестве классификации. Наконец, в выборку может попасть некоторое количество шумовых образов, находящихся области локализации другого класса. Как правило, их удаление только улучшает качество классификации. Эти соображения приводят к идее исключить из выборки шумовые и малоинформативные объекты, оставив только минимальное достаточное количество эталонов.

Рассмотрим вариант решения данной задачи на основе алгоритма STOLP, представленный в. Его основная идея состоит в том, чтобы оставить в качестве прецедентов из каждой обучающей выборки $X^{N_i} = \{x^{(i,1)},\dots,x^{(i,N_i)}\}$ только такие «опорные точки» класса ω_i, которые обеспечивают выполнение следующего условия: расстояние от любой точки обучающей выборки до ближайшего своего образа меньше расстояния до ближайшего прецедента чужого класса. Такой набор прецедентов в качестве эталонов обеспечит безошибочное распознавание всех реализаций обучающей выборки. Контрольные реализации распознаются по методу ближайшего соседа.

Сложность решения этой задачи состоит в том, что состав эталонов прецедентов класса ω_i зависит от того, какие реализации других классов выбраны в качестве прецедентов. Из этого становится очевидным комбинаторный характер задачи, оптимальное решение которой требует полного перебора всех вариантов. Число возможных вариантов оставления по одному прецеденту в качестве эталона для каждого класса равно $\prod_{i=1}^{M} N_i$. Если же оставлять по L прецедентов на класс в качестве эталонов, то число вариантов возрастает до величины $\prod_{i=1}^{M} C_{N_i}^{L}$.
Перебор вариантов можно сократить с помощью алгоритма STOLP – алгоритма отбора эталонных объектов для метрического классификатора. Рассмотрим его работу на примере с двумя классами образов. Более продвинутый вариант алгоритма отбора эталонов называется FRiS–STOLP. При этом используются меры схожести двух объектов относительно некоторого третьего объекта, так называемые FRiS – функции (Function of Rival Similarity). В отличие от упомянутых ранее классических функций расстояния, эта функция позволяет не просто сказать, похожи объекты друг на друга или нет, но и уточнить ответ на вопрос «по сравнению с чем?».

В заключение данной лекции необходимо обратить внимание на задачу распознавания образов, описываемых бинарными признаками, и ее решение с использованием функций расстояния и эталонных образов. Каждый анализируемый образ или эталон представлен в этом случае вектором $x = (x_1,\dots,x_n)^T, x_k \in \{0,1\}$, $z = (z_1,\dots,z_n)^T, z_k \in \{0,1\}$. В подобных задачах обычно используются специальные метрики, такие как нормированное скалярное произведение или так называемая мера Танимото

$$d(x,z) = \frac{x^T z}{\sqrt{(x^T x)(z^T z)}}, \quad d(x,z) = \frac{x^T z}{x^T x + z^T z - x^T z}$$

Структура решающих правил с одним или несколькими эталонами в каждом классе здесь будет аналогична ранее рассмотренной для признаков, являющихся вещественными переменными.

◄ Лекция 7

Перейти на...

Перейти на...

Лекция 9 ►