Лекция 17

Разделение в диагностическом пространстве

Линейное разделение областей диагнозов в пространстве информативных параметров невозможно, если не выполнено условие теоремы о линейном разделении. Построение кусочно-линейных разделяющих функций может оказаться трудоемкой процедурой. В этом случае задача постановки диагноза может быть решена преобразованием исходного пространства параметров в новое *диагностическое пространство*, в котором возможно построение линейных разделяющих функций.

Рассмотрим случай дихотомии. Если области диагнозов в исходном пространстве информативных параметров не пересекаются, то согласно (Л15.6), лекция 15, разделяющую функцию можно представить в виде

$$f(\mathbf{x}) = \sum_{j=0}^{m} c_j \varphi_j(\mathbf{x}), \tag{Л17.1}$$

где индексы i и k опущены. Введем новые координаты, определяемые системой соотношений

$$z_j = \varphi_j\left(\mathbf{x}\right),\ \ j = 1, \ldots, m. \tag{Л17.2}$$

Совокупность координат z_j образует новое m-мерное диагностическое пространство параметров, в котором состояние объекта задается вектором $\mathbf{z} = \left(z_1, z_2, \ldots, z_m\right)$. Преобразование (Л17.2) однозначно преобразует (отражает) исходное пространство в диагностическое. Обратное преобразование в общем случае может быть неоднозначным, то есть одному и тому же значению вектора $\mathbf{z}$ могут соответствовать несколько векторов $\mathbf{x}$.

В диагностическом пространстве разделяющая функция записывается в виде

$$f\left(\mathbf{z}\right) = c_0 + \sum_{j=1}^{m} c_j z_j = 0.$$

Если ввести дополненное диагностическое пространство, добавив координату $z_{m+1} \equiv 1$, и переобозначить коэффициенты $c_j = \lambda_j,\ \left(j = 1, \ldots, m\right);\ c_0 = \lambda_{m+1}$, то разделяющую функцию можно представить в виде скалярного произведения $\mathbf{z} = \left(z_1, z_2, \ldots, z_{m+1}\right)$ и $\boldsymbol{\lambda} = \left(\lambda_1, \ldots, \lambda_{m+1}\right)$, то есть в виде соотношения (Л14.9), см. лекцию 14. Уравнение

$$f(\mathbf{z}) = \boldsymbol{\lambda}\mathbf{z} = 0$$

определяет в дополненном диагностическом пространстве разделяющую гиперплоскость, проходящую через начало координат и нормальную весовому вектору $\boldsymbol{\lambda}$. В итоге задача диагностирования сводится к определению весового вектора $\boldsymbol{\lambda}$, который может быть найден с помощью обучающей выборки объектов и правила постоянного приращения (Л14.15), см. лекцию 14.

Основная трудность при построении диагностического пространства — выбор системы базовых функций $\{\varphi_j(\mathbf{x})\}$. Она должна обеспечивать представление достаточно широкого класса функций в виде ряда, поскольку вид самой разделяющей функции заранее неизвестен. Число членов ряда (Л17.1) может быть большим и в пределе стремиться к бесконечности, то есть диагностическое пространство становится бесконечномерным. С другой стороны, при удачном выборе функций $\{\varphi_j(\mathbf{x})\}$ можно ограничиться небольшой размерностью диагностического пространства, что существенно упрощает вычисления.

Задача. Л17.1. Требуется найти разделяющую функцию двух состояний D_0 и D_1, которые заданы обучающими выборками из 14 объектов с векторами информативных параметров $\mathbf{x} = (x_1; x_2)$, указанными во втором и четвертом столбцах табл. Л17.1 и показанными на рис. Л17.1*а*.

В исходном пространстве параметров условие теоремы о линейном разделении не выполнено. На плоскости (x_1, x_2) отсутствует направление, проекции областей диагнозов на которое не пересекались бы. Разделяющая функция — нелинейная. Введем диагностическое пространство (z_1, z_2). Вид областей D_0 и D_1 (рис. Л17.1*а*) наводит на мысль, что зависимость координат диагностического пространства от параметра x_1 должна быть квадратичная.

Таблица Л17.1. **Векторы признаков обучающей выборки объектов в исходном и диагностическом пространствах**

n	Область диагнозов			
	D_0		D_1	
	Пространство признаков			
	исходное	диагностич.	исходное	диагностич.
	$\mathbf{x}=(x_1;x_2)$	$\mathbf{z}=(z_1;z_2)$	$\mathbf{x}=(x_1;x_2)$	$\mathbf{z}=(z_1;z_2)$
1	(3; 4)	(4; –1)	(2; 6)	(8,5; –4)
2	(4; 3)	(1,5; –1)	(2; 9)	(11,5; –7)
3	(4; 6)	(4,5; –2)	(4; 9)	(7,5; –5)
4	(5; 4)	(2; 1)	(4; 11)	(9,5; –7)
5	(5; 5)	(3; 0)	(6; 10)	(8,5; –4)
6	(5; 6)	(4; –1)	(7; 9)	(9; –2)
7	(7; 4)	(4; 3)	(8; 7)	(9,5; 1)

зададим сравнительно простую зависимость между координатами $(x_1,\, x_2)$ и $(z_1,\, z_2)$

$$\begin{cases} z_1 = 0,5x_1^2 - 5x_1 + x_2 + 10,5; \\ z_2 = x_1 - x_2. \end{cases} \tag{Л17.3}$$

Числовые коэффициенты выбираются произвольно и корректируются так, чтобы области D_0 и D_1 в диагностическом пространстве были линейно разделимы.

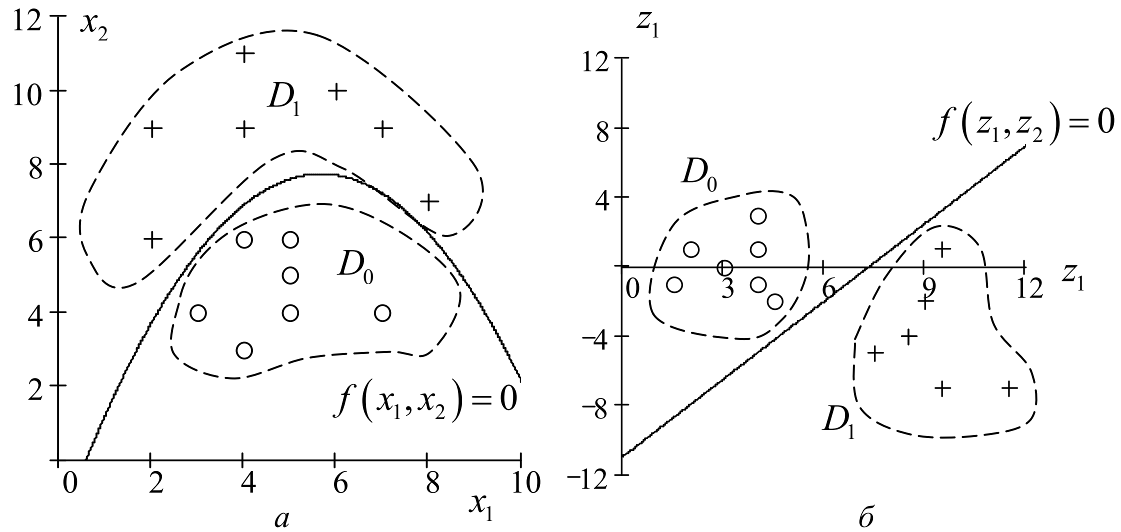

Рис. Л17.1. Расположение векторов обучающих выборок областей D_0 и D_1 в исходном и диагностическом пространствах

Координаты векторов обучающей выборки в диагностическом пространстве, рассчитанные по формулам (Л17.2), указаны в третьем и пятом столбцах табл. Л17.1 и показаны на рис. Л17.1*б*. Построим в диагностическом пространстве линейную разделяющую функцию. Для этого введем дополненное диагностическое пространство с третьей координатой $z_3 \equiv 1$ и сформируем в этом пространстве область $D = D_0 \bigcup \bar{D}_1$, включающую векторы области D_0 и векторы области D_1, у которых изменен знак (табл. Л17.2).

Таблица Л17.2. **Векторы, образующие область** D **в дополненном диагностическом пространстве**

n	$\mathbf{z}_i \in \left(D_0 \bigcup \bar{D}_1\right)$	№	$\mathbf{z}_i \in \left(D_0 \bigcup \bar{D}_1\right)$	№	$\mathbf{z}_i \in \left(D_0 \bigcup \bar{D}_1\right)$
1	(4; –1; 1)	6	(4; –1; 1)	11	(–9,5; 7; –1)
2	(1,5; –1; 1)	7	(4; 3; 1)	12	(–8,5; 4; –1)
3	(4,5; –2; 1)	8	(–8,5; 4; –1)	13	(–9; 2; –1)
4	(2; 1; 1)	9	(–11,5; 7; –1)	14	(–9,5; –1; –1)
5	(3; 0; 1)	10	(–7,5; 5; –1)		

Воспользуемся правилом постоянного приращения, а в качестве первого приближения возьмем весовой вектор, определенный с помощью средних векторов состояний D_0 и D_1 в диагностическом пространстве. В результате вычислений по формуле (Л14.16), см. лекцию 14, находим

$$\mathbf{m}^{(0)} = \left(3{,}29; 0{,}14\right);\ \mathbf{m}^{(1)} = \left(9{,}14; -4{,}00\right);\ \left|\mathbf{m}^{(0)}\right|^2 = 10{,}82;\ \left|\mathbf{m}^{(1)}\right|^2 = 99{,}59.$$

Согласно (Л14.17), см. лекцию 14, весовой вектор в первом приближении имеет следующие координаты $\boldsymbol{\lambda}_1 \approx (-6; 4; 44)$. Непосредственной проверкой можем убедиться, что для всех $\mathbf{z}_i \in D$

$$\boldsymbol{\lambda}_1 \mathbf{z}_i > 0,\ i = 1, \ldots, 14.$$

Следовательно, вектор $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1$ является весовым. Уравнение

$$-3z_1 + 2z_2 + 22z_3 = 0$$

в дополненном диагностическом пространстве определяет разделяющую плоскость, которой в диагностическом пространстве соответствует разделяющая прямая (рис. Л17.1*б*)

$$f\left(z_1, z_2\right) = -3z_1 + 2z_2 + 22 = 0.$$

Подставив в это соотношение z_1 и z_2 из (Л17.3), получим уравнение разделяющей линии областей диагнозов в исходном пространстве параметров

$$f\left(x_1, x_2\right) = 1{,}5x_1^2 - 17x_1 + 5x_2 + 9{,}5 = 0,$$

которое описывает параболу, показанную на рис. 8*а*.