

$$\begin{pmatrix} 1 & X_{11} & \cdots & X_{1n} \\ \vdots & \vdots & & \vdots \\ 1 & X_{k1} & \cdots & X_{kn} \\ -1 & -X_{k+1,1} & \cdots & -X_{k+1,n} \\ \vdots & \vdots & & \vdots \\ -1 & -X_{m1} & \cdots & -X_{mn} \end{pmatrix} \cdot \begin{pmatrix} W_0 \\ W_1 \\ \vdots \\ W_n \end{pmatrix} > \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (1)$$

记作 $X_1 \cdot W^T > 0$

$X_2 \cdot W^T > 0$ 或 $\|X\| \cdot W^T > 0$

.....

$X_m \cdot W^T > 0$ (T表示转置)

矩阵 $\|X\|$ 只是将原样本矩阵中的B类样本的各个特征变量(包括虚设变量)的赋值统统加上负号, $X_1, X_2, \dots, X_m$ 则表示矩阵 $\|X\|$ 中的各个行向量。

实际上解不等式组(1)的方法很多,一般可采用消去法,迭代解法(包括张弛法、指数法图象识别法等)。但应指出,不等式组(1)若有解(即相容),则它有无穷多解,且构成一个开角锥¹⁾,而图象识别法只是求出其中的某一个解(而且前面的反例说明,随着W向量初始值的不同,样本的简化或样本顺序的改变,这个解常常是会随之改变的)然后就以这个解作为权重向量确定判别函数,进行识别。于是自然产生一个疑问,即这无穷多个解所对应的无穷多个判别函数能不能得出相同的预测结果?至少在某些情况下这是不可能的,但究竟在什么情况下会得出不同的预测结果,在什么情况下又会得出一致的预测结果呢?为解决这个问题,我们先要讨论组(1)的相容性。即训练图象的线性可分性。

二、线性不等式组的相容条件

根据线性不等式组的理论^[4]可直接推出

定理1 不等式组(1)相容的充要条件是:不存在不全为零的非负数 $\lambda_1, \dots, \lambda_m$,

$$\text{使 } \sum_{i=1}^m \lambda_i X_i = 0。$$

由定理1立即可推出

定理2 不等式组(1)不相容的充要条件是:存在不全为零的非负数 $\lambda_1, \dots, \lambda_m$,

$$\text{使 } \sum_{i=1}^m \lambda_i X_i = 0。$$

如果设 $\lambda_1, \dots, \lambda_m$ 中大于0的为 $\lambda_{i_1}, \lambda_{i_r}$ ($1 \leq i_1, \dots, i_r \leq m$)则得出

定理2' 不等式组(1)不相容的充要条件是:存在 $\lambda_{i_1} > 0, \dots, \lambda_{i_r} > 0$, ($1 \leq$

$i_1, \dots, i_r \leq m$)使 $\sum_{k=1}^r \lambda_{i_k} X_{i_k} = 0$, 换句话说,即矩阵 $\|X\|$ 中有其几行正系数线性相关。

三、关于预测结果的确定性

在这部分讨论中我们总假定组(1)是相容的,因为在不相容的情况,两类训练图象是

1) 角锥即多面凸锥(详见文献[4]),开角锥即不包含边界面的多面凸锥。

线性不可分的。

我们称向量 $\mathbf{b} = (1, b_1, \dots, b_n)$ 为待预测(识别)向量, 对其一待预测向量 $\mathbf{b}$ 显然只有下列三种可能结果:

- i) 对组(1)的所有解 $\mathbf{W}$, 都有 $\mathbf{W} \cdot \mathbf{b}^T > 0$, 此时对 $\mathbf{b}$ 可作出确定的预测—A类。
- ii) 对组(1)的所有解 $\mathbf{W}$, 都有 $\mathbf{W} \cdot \mathbf{b}^T < 0$, 此时对 $\mathbf{b}$ 可作出确定的预测—B类。
- iii) 除上述两种情况, 都不可能对 $\mathbf{b}$ 作出确定的预测。

下面讨论这三种结果各在什么条件下发生。

定理 3 对组(1)的所有解 $\mathbf{W}$, 都有 $\mathbf{W} \cdot \mathbf{b}^T < 0$ 的充要条件是: $\mathbf{b}$ 可表为矩阵 $\|\mathbf{X}\|$ 中某几行的负系数线性组合。

证明 充分性是明显的。

为证必要性, 讨论下列不等式组:

$$\begin{aligned} \mathbf{X}_1 \cdot \mathbf{W}^T &> 0 \\ \dots\dots\dots &\text{记为 } \|\mathbf{X}'\| \cdot \mathbf{W}^T > 0 \\ \mathbf{X}_m \cdot \mathbf{W}^T &> 0 \\ \mathbf{b} \cdot \mathbf{W}^T &> 0 \end{aligned} \quad (2)$$

组(2)只是比组(1)增加了一个不等式 $\mathbf{b} \cdot \mathbf{W}^T > 0$ 。

若对(1)的所有解 $\mathbf{W}$ 都有 $\mathbf{b} \cdot \mathbf{W}^T < 0$, 则(2)显然不相容, 由不相容的充要条件(见定理2')可知, $\|\mathbf{X}'\|$ 中某几行正系数线性相关, 由于(1)是相容的, 所以其中必包含行向量 $\mathbf{b}$, 设

$$\alpha_{i1}\mathbf{X}_{i1} + \dots + \alpha_{ip}\mathbf{X}_{ip} + \beta\mathbf{b} = 0 \quad (\text{式中 } \alpha_{i1}, \dots, \alpha_{ip}, \beta \text{ 皆大于 } 0)$$

$$\text{则 } \mathbf{b} = -\frac{\alpha_{i1}}{\beta}\mathbf{X}_{i1} - \dots - \frac{\alpha_{ip}}{\beta}\mathbf{X}_{ip}。 \text{故必要性获证。}$$

同理, 对不等式组

$$\begin{aligned} \mathbf{X}_1 \cdot \mathbf{W}^T &> 0 \\ \dots\dots\dots \\ \mathbf{X}_m \cdot \mathbf{W}^T &> 0 \\ -\mathbf{b} \cdot \mathbf{W}^T &> 0 \end{aligned}$$

进行完全类似的讨论可得

定理 4 对(1)的所有解 $\mathbf{W}$ 都有 $\mathbf{b} \cdot \mathbf{W}^T > 0$ (即 $-\mathbf{b} \cdot \mathbf{W}^T < 0$) 的充要条件是: $-\mathbf{b}$ 可表为 $\|\mathbf{X}\|$ 中某几行的负系数线性组合, 即 $\mathbf{b}$ 可表为 $\|\mathbf{X}\|$ 中某几行的正系数线性组合。

至于结果iii)是否总有可能发生, 我们有下面的结论(证明较繁, 见附录)

定理 5 若组(1)相容, 则至少存在二个开角锥 P 与 P^* , 当 $\mathbf{b} \in P$ (或 P^*) 时, 对 $\mathbf{b}$ 不能作出确定的预测。

定理5说明对于图象识别法来说, 产生预测结果的不确定性不是偶然的。

至此, 问题已从理论上完全解决。

四、小结

由定理1—5, 我们得到下述结论:

1. 若矩阵 $\|\mathbf{X}\|$ 中某几行正系数线性相关, 则组(1)不相容, 否则组(1)是相容的,

2.若组(1)相容,且待预测向量 b 可表为 $\|X\|$ 中某几行的正系数线性组合时,可对 b 作出确定的预测—A类;

3.若组(1)相容,且 b 可表为 $\|X\|$ 中某几行的负系数线性组合时,可对 b 作出确定的预测—B类;

4.若组(1)相容,且 b 不属于上述2、3二种情况则对 b 不能作出确定的预测。

可见,用图象识别法进行分类预测,在预测结果确定的情况是可行的,算的结果与上述结论2、3是一致的,而在预测结果不确定的情况(即结论4所述的情况),用图象识别法随机算出一个结果是不合适的,最好换用别的方法。

但图象识别法还是可以考虑改进的,关键是增添条件使解唯一确定,如果增添使解的某个线性(或非线性)函数达到极大(或极小)的条件,则可变为线性(或非线性)规划的问题。这是一个设想,究竟如何改进比较合理有效,还有待进一步研究解决。

本文的定理和结论都可用多面凸锥(即角锥)的术语进行叙述和论证,但为了避免繁琐的数学定义,我们直接采用“正(负)系数线性相关(组合)”等语言叙述。

(本文1981年11月21日收到)

PROBLEMS AND DISCUSSIONS ON THE APPLICATION OF PATTERN RECOGNITION IN ZONING STRONG MOTION AREA

Mao Weimin

(The Teachers Training School of Baotou)

Abstract

There are two problems in zoning the strong motion area by means of pattern recognition. One is how to decide whether the training set is divisible linearly, which is directly related to the problem of convergence of the iteration process. The other problem is the conflicting results of prediction that come from the differences of recognition functions which are brought about by the fact that the solution of training is not unique. In this paper these two problems are discussed according to the theory of system of linearly inequalities and a preliminary result has been obtained.