子空间法模式识别

〔芬兰〕E. 奥亚 著

蔡国廉 杨文瑜 译

科学出版社

内 容 简 介

本书系统地介绍了子空间法模式识别的理论基础、数值算法和子空间分类器设计等内容. 全书共分五章. 第一、二两章主要叙述子空间法模式识别的特点和所需的数学预备知识，第三章为 K-L 变换的数值计算，第四章为子空间分类器的设计，第五章为学习子空间算法及其在语音识别中的应用. 书末附有随机近似和收敛证明两个附录以及本学科领域的最新参考文献. 本书是一本较有实用价值的参考书.

本书的特点是：内容新颖，概念清晰，叙述简明扼要，深入浅出，便于自学.

本书可供从事计算机、自动化技术、信息处理、生物医学工程以及模式识别应用领域中的科技人员和高等院校有关专业师生参考.

E. Oja
SUBSPACE METHODS OF
PATTERN RECOGNITION
Research Studies Press Ltd., 1983

子空间法模式识别

〔芬兰〕E. 奥亚 著
蔡国廉　杨文瑜 译

责任编辑　鞠丽娜

科学出版社出版
北京朝阳门内大街 137 号

中国科学院印刷厂印刷

新华书店北京发行所发行　各地新华书店经售

*

1987 年 9 月第一版　开本：787×1092　1/32
1987 年 9 月第一次印刷　印张：5 3/8
印数：0001—8,200　字数：116,000

统一书号：15031·859
本社书号：5403·15—8

定价：1.30 元

译 者 的 话

模式识别技术是近二十多年来发展极为迅速的一门新学科．传统上通常将模式识别方法概括为两大类，即统计方法和句法方法．所谓模式的分类，实质上就是在特征空间中划区，不同的区域对应不同的类别．而子空间法模式识别的特点,则是把模式的类别由一些线性子空间来限定,每一个线性子空间对应一个类别．每类只把统计上最为有用的特征保留用来张成子空间,各类子空间的维数也可互不相同．一个矢量对某类子空间的投影，即作为衡量该矢量与该类模式之间的类似度．因而子空间法判别函数的分类只需计算为数不多的几个矢量内积，就可使运算速度大为加快．本书讨论了子空间法模式识别的理论基础，重点在于 K-L 变换等统计正交展开的数值计算，所介绍的基于样本协方差矩阵的方法和从样本直接计算的各种递归算法都具有特色，其中 K-L 变换的快速算法更为一般书籍中所鲜见．在子空间分类器的设计中，分别讨论了没有学习和有学习的子空间分类算法，并以语音识别中频谱的分类为实例加以讨论，此部分内容可作为子空间法在其它领域中应用的借鉴．

本书是英国 Research Studies Press Ltd. 出版的"Pattern Recognition and Image Processing Series"(模式识别与图象处理丛书)的第六本．作者 E. 奥亚博士为芬兰 Kuopio 大学应用数学系副教授，近年来在"模式识别"等国际性杂志上和有关国际会议上发表过多篇有关子空间分类算法的论文，本书不仅系统地总结了该学科的最新发展动态，同时也是作者在

研究语音识别中应用子空间法对音素进行自动分类的科研成果．我们选译本书的目的则是希望它不仅对从事信息处理和模式识别方面的专业人员有所帮助，而且对子空间法模式识别在我国的深入研究，也能起到促进作用．

本书的序、第一、二、五章和附录1由蔡国廉翻译，第三、四章和附录2由杨文瑜翻译．上海交通大学李介谷教授对全书译文作了认真的校阅，提出了宝贵意见，特此致谢．

由于译者水平有限，译文中难免有不妥之处，敬请读者批评指正。

译者

1986年4月

序

本书的主题

在模式识别、特征抽取、以及通常所说的信号处理领域内使用矢量空间法时，线性变换起着举足轻重的作用. 应用线性变换技术，可使模式表达更为紧凑，而且在正确分类中起重要作用的一些性质也被突出出来. 通常线性变换可以理解为：把矢量表达的信号或其模式，由其原始表达的空间映射到低维的子空间中去.

作为上述基本过程的一种开拓，子空间法模式识别应运而生. 此法的技巧在于，模式的类别不是预先规定为特征空间中的一些有限范围或区域，而是由一些线性子空间来限定，每一个线性子空间对应于一个类别. 一个矢量和某类别子空间之间的距离，作为衡量该矢量与该类模式之间的类似度或隶属度，同时它提供了子空间法的判别函数.

子空间法最宜用于这样的数据矢量，其中矢量各分量之间的相对强度或相对能量比起信号的总的电平来说更加重要得多. 子空间法能使高维矢量表达紧凑，分类速度加快. 每类只把统计上最有用的特征留在子空间表达中. 分类规则也只要计算为数不多的几个矢量内积，因而颇为简便.

本书讨论了子空间法的理论基础及所采用的各种研究途径，着重于在自动语音识别及更为一般的频谱分类中使用的学习子空间法. 子空间法在模式识别和图象处理的各种领域中已获得广泛的应用. 对子空间法的数学基础及其应用，本书均作了全面的阐述.

关于数学预备知识，统计正交展开诸如主要分量法和

K-L(Karhunen-Loève) 展开及其在信号处理中的应用等均被述及．这类展开，对之进行定义往往并不困难，而要进行实际运算却很不容易，因此本书中介绍了几种求取基本矢量的数值算法．这些算法都涉及到对于多类模式识别设计最佳类别子空间的学习过程．

在频谱矢量分类试验中，将子空间分类器获得的结果与某些标准分类器的性能进行了比较，其中包括在识别一千个词汇量的单词时对音素数据的广泛测试．结果表明，这种利用子空间对模式进行识别的方法，至少对某些通用的数据类型来说是颇具诱惑力的．

感谢

本书大部分的素材是作者在 Helsinki 技术大学计算机与信息科学研究室工作期间的研究成果． 研究室主任 T. Kohonen 教授无论在课题的指导思想和全面的组织领导中，都作出了杰出的贡献．衷心感谢他首先引导我进入子空间分类器的课题研究，并对其中许多重要的数学问题进行了反复讨论．本书所举子空间法的主要应用为音素自动识别，这是该研究室已着手进行的语音识别研究规划的一部分，参加的成员有 S. Haltsonen 博士、M. Jalanko 博士、E. Reuhkala 博士和 H.Riittinen 先生．我非常感谢他们允许我在写作本书时使用他们的研究成果，并在计算机试验过程中提供了各种现成的试验资料．对 J. Karhunen 先生和 M. Kuusela 女士，感谢他们编译了大量的音素数据试验的计算机程序，并付诸运行，同时还进行过多次的讨论．K. Eloranta 先生和 J. Parkkinen 先生也在许多方面对我进行过帮助．本书稿由 I. Halenius 女士熟练地打印完成．

芬兰科学院和 Niilo Helander 基金对本研究课题在财政上的有力支持，使这本著作得以问世，在此表示万分感谢．

常用符号表

（依字母排列）

符号	定义
A_k	随机矩阵
$c_i\ (c_i^{(i)})$	协方差矩阵的第 i 个本征矢量（第 i 类的第 i 个本征矢量）
$\delta(.)(\delta^{(i)}(.))$	判别函数（第 i 类的判别函数）
$\delta(.,.)$	距离函数
δ_{ij}	克罗内克 (Kronecker) δ
$E(.)$	期望值
$E(.\|.)$	条件期望值
γ_k	在递归估计中的增益参数
k	常用迭代标志
K	类别数
$L(L^{(i)})$	子空间（第 i 类子空间）
$L(a_1,\cdots,a_p)$	由矢量 $a_1,\cdots,a_p$ 张成的子空间
$\lambda_i\ (\lambda_i^{(i)})$	协方差矩阵的第 i 个本征值（第 i 类的第 i 个本征值）
$\mu, \mu_o, \mu_r, \mu_k^{(i,j)}$	LSM 法中的旋转参数
n	模式空间的维数
$p\ (p^{(i)})$	子空间的维数（第 i 类子空间的维数）
$P\ (P^{(i)})$	投影矩阵（第 i 类投影矩阵）
$P(.)$	概率
$P(.\|.)$	条件概率

$\pi^{(i)}$	第 i 类的先验概率
$Q\ (Q^{(i)})$	相关矩阵(第 i 类的相关矩阵)
$r\ (r^{(i)})$	样本的大小 (第 i 类的样本大小)
$R\ (R^{(i)})$	协方差矩阵 (第 i 类的协方差矩阵)
R^n	n 维实数空间
$R^{n\times m}$	$n\times m$ 实数矩阵空间
$\sigma_k^{(i)}$	第 i 个本征值的递归估计
$T\ (x^T)$	转置 (矢量 x 的转置)
τ_k	矩阵迹的递归估计
$u_k^{(i)}$	第 i 个本征矢量的递归估计
$\omega^{(i)}$	第 i 类
x	模式矢量
x_i	样本矢量
$\hat{x}$	x 对子空间的投影
$\tilde{x}$	垂直残差
ξ_i	矢量 x 的第 i 个分量
z_i	具有零平均的样本矢量

目　　录

第一章 引 言

传统上，称之为决策理论（几何或矢量空间）模式识别的课题由两个彼此无关的步序所组成，即是：特征选择或特征抽取和模式分类，然而，在近代技术水平下，这种简单而又有些任意的划分方法，已被较为复杂的观点所替代．事实上，在大多数方法中都含有对数据的结构分析，它可以包括建立模型、特征抽取、聚类和统计检验等诸项．还有，任何一种模式识别方案的第一步都是从通常极为大量的粗数据中把对象表达出来，这时某种数据压缩是必不可少的． 在这种情况下，使信息损失最小是一个中心的目标．一般来说，任何一个信号、波形或图象都含有不少冗余信息，这些冗余信息可以采用例如级数展开等方法加以去除．

很多这种预处理、特征抽取和数据压缩技术都可以在数学上表达成线性变换的形式．由于总是希望把原始测量数据集用更紧凑的方式来表达，因此这种变换都是对低维子空间的一种线性映射．

一个采样波形的傅里叶系数子集可作为一例，若把傅里叶系数当做特征，则挑选子集就相当于选择特征．用几何术语来说，经这样的选择后，原始特征空间中的每项数据都用其线性子空间来表达．

某些统计数据压缩和分析技术，譬如 K-L 变换和主要分量分析等，也能产生低维的子空间．它们只是一些简单的线性展开，其中的基本矢量数目比构成整个模式或特征空间所要求的矢量数目有所减少．

在较为一般的信息处理理论中，已经把子空间用来表达结构信息．在联想信息处理的理论中（参见 Kohonen (1977, 1983), Kohonen 等(1981)的著作及其中的参考文献)，投影原理和子空间的使用，在本质上说明了联想映射的最优化性质．Kohonen 指出，实现这类诸如转移算子的映射的物理网络，其工作可能象高度平行的分布式联想存贮器一样．1977 年 Kohonen 等对联想算子和有关的子空间的模式处理提出了某些说明．

Watanabe 在 1969 年的论文中讨论了子空间与逻辑之间的关系．他引入了一个逻辑命题的模格 (modular lattice)，其中每个子空间被赋以一个命题．Grenander 于 1976 年和 1978 年也在模式理论中提出了类似的概念，其中一个抽象模式分析器和模式处理器的逻辑可用希耳伯特空间中投影算子的代数分析来给定．

从逻辑上考虑，Watanabe 建议子空间法模式识别可作为分类和表达模式的一种新方法，这时模式成为矢量空间中的元素．子空间法把每类中最显著的特征个别地抽取出来，并用一个模板集来加以表达．其中心思想是每类各有自己带代表性的特征集，类与类之间的特征集各不相同，特征数目也可以彼此相差很大．Watanabe 提议各类的基本特征集可从 K-L 展开中求得．

从几何角度考虑，子空间法意味着多类模式识别问题中的每一类对应于原始线性模式空间中的一个子空间．这些子空间分别由每一类的特征矢量构成的矢量级数展开所张成．这种子空间法对模式识别的传统概念也很吻合：一方面，子空间法能将一个输入样本分到最为合适的类别中去；另一方面，子空间表达式中也提示了该类总体结构中的某些性质．由于每一类别的基本结构信息都能在表达该类的子空间中得

到反映，因此从这个意义上看，子空间法和结构法模式识别具有共性．

对于所提出的任何一种模式识别方案的最后检验则是它在实际中的工作如何．这不仅意味着要有很小的识别误差率，同时还要求对分类规则计算能达到很高的速度．诚然，分类器和量测机构之间的接口以及后处理系统都是不容忽视的重要因素．在这些方面，正是子空间法所具有的优点．由于它将通常的特征抽取和模式分类两步合并在一步内完成，所以能处理含有许多分量的原始数据矢量．分类规则只要进行少量的矢量内积运算，就可使计算速度很快，不过其决策区域仍然具有二次型的边界．子空间分类器通过合理的设计，可达甚低的误识率．分类器的输出，不仅含有必须的类别隶属函数，而且还包括了矢量在类别子空间上投影后对已变换数据的描述．

概括起来，本书考虑了两个中心课题．一是用于数据压缩的信号子空间，这时只需一个子空间；另一种是用于分类的子空间，此时要用若干个子空间，每类一个．模式识别与信号处理中使用的上述两种子空间并非彼此无关．子空间分类器的起点是对每类分别进行线性特征抽取，因而一个类别的情况成为需求解的一个重要子问题．此外，Kohonen（参阅 Kohonen 等（1978）的论文）介绍了用于多类别情况下的学习子空间法，也提出了一种在信号空间中只有一个低维子空间时的新型迭代算法，它使用了 K-L 展开．

在书中始终处于主导地位的内容将成为我们在两方面的研究内容：学习子空间法以及与之有关的由统计正交展开计算基本矢量的递归算法．书中列举了一些有关工作的参考文献及应用实例，以供对此感兴趣的读者参考．

第二章为基本数学工具，目的在于理解子空间分类器和

有关学习规则的性能，其中也引用了线性代数和 K-L 展开等正交统计展开的研究成果方面的内容.在具体应用展开式时，由于受运算速度和存贮器容量所限，不能完全直接地对它们进行实际计算. 第三章讨论了求取展开式基本矢量的几种计算方法. 它包括基于样本协方差矩阵和相关矩阵的一些矩阵代数中常规的算法和有关随机矩阵的本征矢量递归计算的一些研究新成果. 近年来在信号处理应用方面采用了这些新成果，而其理论结果对以后的子空间分类器学习算法来说也是必不可少的. 第三章还详细讨论了递归算法，它是数学概念上称之为“子空间迭代”的一个应用实例. 这也是学习子空间法的关键技术所在.

第四章综述了非学习子空间法. 内容有代数法、统计法及其推广应用于紧缩和正交子空间法. 对基本统计模型的推广进行了某些评论. 第五章综述和分析了 Kohonen 学习子空间法，阐述了它在自动语音识别中频谱音素分类和在多光谱卫星图片象素分类中的应用. 给出了某些理论上的收敛性结果，同时推导出作为 Kohonen 法变种的平均学习子空间法，并应用于语音识别中.

第二章　子空间与统计正交展开

子空间法模式识别建立在线性正交展开的基本原理上，根据正交基本矢量或正交函数进行的这种正交展开是矢量空间法的一种主要的工具．本章2.1节对线性代数的某些重要数学概念作了简要的回顾，其后在2.2节对统计正交展开进行了讨论．

2.1　子空间法的数学预备知识

子空间法实质上是沿袭了希耳伯特空间方法论中的诸如内积、正交基和投影算子等基本概念．但是，实际上模式空间的维数是有限的，因而如果将数学背景仅限于欧氏空间和普通的实数矩阵代数，则不会有所失真．全书有一些统一的符号规定，所采用的符号与数字矩阵代数标准文本的符号颇为接近．

n 维实数欧氏空间记为 R^n．用小写罗马字母表示矢量，标量则大部分用小写希腊字母表示．对作为指数的整数、多个变量之和等此规约有一个例外．例如

$$x = (\xi_1, \xi_2, \cdots, \xi_n)^T$$

表示矢量，其分量为实数标量 $\xi_1, \cdots, \xi_n$．矩阵全部用大写的罗马字母和希腊字母表示，如 A，Γ 等．矩阵用其列矢量分划时记为

$$A = (a_1 a_2 \cdots a_m)$$

$n \times m$ 的矩阵空间记为 $R^{n\times m}$．

本书常用的一些变量和数量的专用符号列于卷首.

x^Ty 和 $(x^Tx)^{1/2}=\|x\|$分别表示内积和模. 基是一组线性无关的矢量,比如说记为 $a_1,\cdots,a_p$. 该矢量集通常总是正交化的.下面介绍的格拉姆-施密特正交化(GSO)(Gram-Schmidt Orthogonalization) 过程是一种很实用的算法, 当然它并不是唯一的方法.

$$b_1=a_1$$

$$b_j=a_j-\sum_{k=1}^{j-1}\frac{a_j^Tb_k}{b_k^Tb_k}b_k,\ j=2,\cdots,p \tag{2.1}$$

通常, 当 $i\neq j$ 时 $b_i^Tb_j=0$. 如果进一步将每个 b_j 用其模去除,则这个集将是正交归一的: $b_i^Tb_j=\delta_{ij}$. 有关 GSO 和其它有关的数值计算方法的参考文献可查阅 Householder (1964) 的著作.

用一组线性无关的矢量集 $\{a_1,\cdots,a_p\}$ 所张成的 R^n 的子空间记为 $L(a_1,\cdots,a_p)$,并定义为

$$L=L(a_1,\cdots,a_p)=\{z|z=\sum_{i=1}^{p}\zeta_ia_i,$$

$$\text{对标量 }\zeta_1,\cdots,\zeta_p\} \tag{2.2}$$

张成子空间的线性无关矢量的数目 p 为 L 的维数, 记作 $p=\dim(L)$. 在计算机中用数值形式表达一个子空间的最方便的方法是取其正交归一的基本矢量, 它唯一地确定了一个子空间,见图 2.1.

一个仿射子空间或线性流形的形式为

$$M=\{z|z=x_0+x_1,\text{其中 }x_1\in L,x_0\text{ 固定}\} \tag{2.3}$$

若 $x_0\neq 0$,则 R^n 的原点不在M上, 而任意一个子空间 L 中总是包含有原点. 式 (2.3) 也可写为 $M=x_0+L$. 图 2.2 画出了一个仿射子空间的例子.

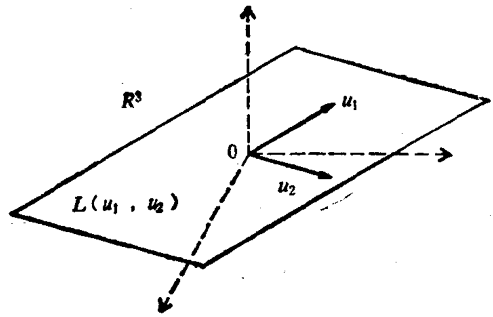

图 2.1 在 R^3 中的一个二维子空间．矢量 u_1, u_2 为正交归一基本矢量

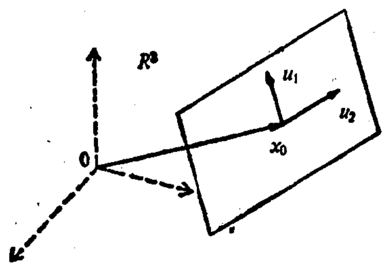

图 2.2 R^3 中的仿射子空间 $x_0 + L(u_1, u_2)$

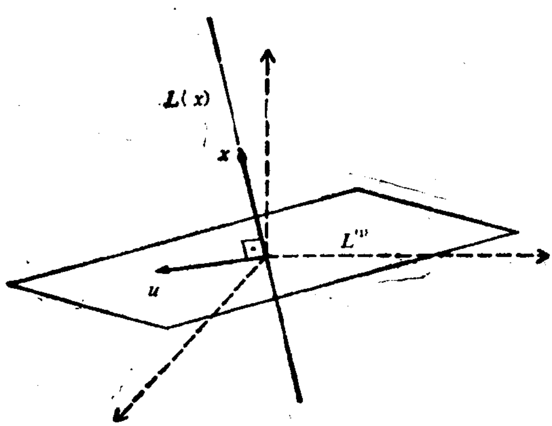

图 2.3 矢量 x 垂直于 $L^{(1)}$，因此也垂直于任何 $u \in L^{(1)}$ 的矢量．而且子空间 $L(x) \perp L^{(1)}$

矢量 x 与子空间 L 具有正交性，记为 $x \perp L$，对于一切 $u \in L$ 可根据内积 $x^T u = 0$ 来加以定义. 两个子空间 $L^{(1)}$ 与 $L^{(2)}$ 具有正交性，即意味着对于一切 $u \in L^{(1)}$，$v \in L^{(2)}$，$u^T v = 0$（图 2.3）.

子空间法的一个基本概念为矢量 x 在子空间 L 上的投影. 假设由一组正交归一的基本矢量 $\{u_1, \cdots, u_p\}$ 张成子空间 L，并且满足 $u_i^T u_j = \delta_{ij}$. 则矢量

$$\hat{x} = \sum_{i=1}^{p} (x^T u_i) u_i \tag{2.4}$$

为 x 在 L 上的（正交）投影. 它有两个重要性质：第一，矢量 $x - \hat{x}$ 垂直于 L；第二，$\hat{x}$ 为 L 中使 $\|x - y\|$ $(y \in L)$ 最小的一个唯一的矢量（证明参见 Albert (1972) 的论文）. 矢量 $x - \hat{x}$ 为垂直残差，记为 $\tilde{x}$. 图 2.4 说明投影与残差矢量之间的关系. 对某一给定的子空间 L，我们经常将矢量分解为

$$x = \hat{x} + \tilde{x} \tag{2.5}$$

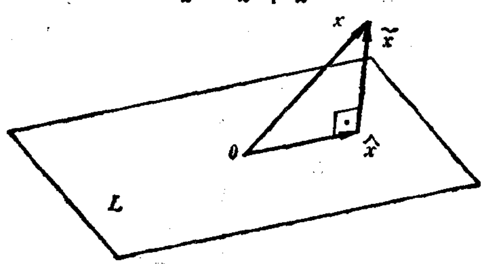

图 2.4 矢量 x 在 L 上的投影与垂直残差

若 L 为一个子空间，并设 $x_0 \neq 0$，则矢量 x 在线性流形 $M = x_0 + L$ 上的（正交）投影为 $x_0 + \hat{\bar{x}}$，其中 $\hat{\bar{x}}$ 为矢量 $x - x_0$ 在 L 上的投影（图 2.5）.

实用上，投影矢量 $\hat{x}$ 应由式 (2.4) 计算. 在模式识别应用中，子空间的维数 p 往往比模式空间的维数 n 少得多. 因此若使用专用的处理器，其矢量内积能计算得非常快. 事实