

李群机器学习研究综述

李凡长 何书萍 钱旭培

(苏州大学计算机科学与技术学院机器学习与数据分析研究中心 苏州 215006)

摘 要 文中简述了李群机器学习的相关研究内容, 包括李群机器学习的概念、公理假设、代数学习模型、几何学习模型、Dynkin图的几何学习算法、量子群、辛群分类器的设计、轨道生成学习算法等.

关键词 李群机器学习; 公理假设; 李群; 分类器

中图法分类号 TP18 DOI号: 10.3724/SP.J.1016.2010.01115

Survey on Lie Group Machine Learning

LI Fan-Zhang HE Shu-Ping QIAN Xu-Pei

(School of Computer Science and Technology, Machine Learning and Data Analysis Research Center, Soochow University, Suzhou 215006)

Abstract This paper summarizes the relevant research of Lie group machine learning, including: conceptions of Lie group machine learning, assumption axioms, algebra learning model, geometric learning model, geometric learning algorithms of Dynkin diagram, learning algorithms of orbits generated and so on. Especially, this paper gives out specific design of quantum group classifier and symplectic group classifier.

Keywords Lie group machine learning; assumption axiom; Lie group; classifier

1 引 言

李群机器学习(Lie Group Machine Learning, LML)作为机器学习领域的一种新的学习方法,一方面继承流形学习的优点,另一方面借用李群的思想,形成了具有创新特色的学习范式.自2004年提出至今^[1-47],已引起加拿大、爱尔兰、芬兰、意大利、美国等国内外同行的广泛关注.该方法和流形学习方法相比有明显优势,从李群的概念可以看出,它包含了微分流形和群的内容;微分流形包含了拓扑流形和微分结构.这套理论系统,既给我们提供了描述数据的几何表示方法,又给出了具体的代数求解方案.如:群保持了系统的完备性,微分提供了具体的代数计算方法,流形给出了几何表示方法,这正符合

认知理论中的定性和定量表示相结合的认知模式.另外,从认知过程来看,人脑在认识客观世界中的任何对象时,首先关注的是表征问题的一个稳定点,然后依次进行分析图像的结构.对这种认知模式,在李群结构中,最小生成元就是这个稳定点,只要找到这个最小生成元,就可以利用李群方法对图像进行了分析.因此,李群机器学习既符合学习认知规律,又满足计算机解决现实问题的条件.

文献[31]提出了基于李群理论的一个贝叶斯方法来学习视觉感知,运用基于指数矩阵的图形生成模型,从包含微小变换的输入数据得到一个学习李群算子的非监督学习算法.文献[32]使用运动表示的内在李群结构来求平均,用特殊正交群 $SO(3)$ 和特殊欧几里得群 $SE(3)$ 的李代数来定义李群上的平均值.提出了全局相容运动估计的李代数求均值法,

此方法能够线性计算所有可能的相对运动以及对运动估计快速求平均. 文献[33]提出了一种新的二维轮廓的不变特征不变信号. 它是对在从李群变换群中得到的许多变换下的轮廓不变程度的度量, 其中李群理论提供了在一个变换动作下, 点位置的局部变换和此变换的全局描述的一种联系, 并提供了自然的起始点, 说明了不变信号在转移、旋转和轮廓缩放后本身保持不变. 文献[34]用李群方法来寻找特征空间中的健壮和稳定特性, 提出了基于李群理论的支配子空间不变量的概念和拟不变量的一种特殊类型, 并给出了支配子空间不变量(DSI)算法. 文献[35]给出了一种由有噪声的二维图像流对三维刚体运动进行神经计算的新方法. 将对三维图像流解释看成是一个线性信号变换, 由神经网络来执行, 其基本信号是三维欧几里得群的 6 个极小生成子的二维矢量场. 神经网络模拟结果表明, 在随机噪声导致传统的代数方法失败时, 此方法仍能保证其可靠性. 文献[36]给神经系统的权空间赋予了一个特殊李群结构, 用李群的微分几何结构来学习这些属性并揭露不同学习方法间内在属性的联系, 从而提出了基于李群的非监督神经学习. 文献[37]提出了在 Stiefel 流形上基于测地线和接近于测地线的曲线簇(拟测地线)的新的神经网络学习算法, 详细地解释了李群本身的梯度和测地线如何导出其在齐性空间上的对应, 强调了李群而非正规化约束在这个流形上的作用. 文献[38]提出了图像分析中的约束主轮廓进化的李群方法. 在此方法中, 使用平面作用的李群及其极小生成子的李代数的对应关系适当地调整欧拉-拉格朗日下降方法, 以此保证在保持原函数的下降方程不变时, 曲线进化发生在被选变换群的轨道上, 此方法的优点在于无需任何修改就能直接实现曲线进化, 在许多跟踪和分割的应用中起到了重要作用. 文献[39]引入一个能够在分布式控制网络上执行的线性算子类, 称为回归正交变换(ROTS). ROTS 可以表示特殊正交群 $SO(n)$ ($n \geq 4$), 通过对其对应的内在的李群结构中提取梯度流来找到一个适合工作的 ROT. 文献[40-54]提出了新的度量学习问题的框架结构, 将候选度量集限制在一个参数族中, 此度量是在一个李群变换下的模糊的信息量的拖回度量. 文献[41]提出了机器学习中的群论和非交换调和与分析相结合的一种学习问题研究框架, 给出了群作用下核不变量的对称理论、在傅立叶谱下群定向核的转换不变量的特性及置换学习的 Perceptron 算法等相关内容.

本文第 2 节介绍李群机器学习的基本概念; 第 3 节介绍李群机器学习的公理假设; 第 4 节介绍李群机器学习的学习模型; 第 5 节介绍李群机器学习中的 Dynkin 图几何学习算法; 第 6 节介绍李群机器学习的分类器设计; 第 7 节介绍李群机器学习子空间轨道生产学习算法; 第 8 节给出结论与展望.

2 李群机器学习的基本概念

先对李群做些说明.

定义 1. 设 G 是一个非空集合, 满足

- (1) G 是一个群;
- (2) G 也是一个微分流形;
- (3) 群的运算是可微的, 即由 $G \times G$ 到 G 的映射 $(g^1, g^2) \mapsto g^1 g^2$ 是可微的映射.

则称 G 是一个李群(Lie group).

从李群的定义可以看出: 李群既是一个群, 同时也是一个微分流形. 我们知道, 流形是点、线、面以及各种高维连续空间概念的推广, 而我们在用机器学习方法分析数据时, 所有观测数据都是可以和点、线、面等结构建立起对应关系的. 从流形的角度, 文献[1-8]给出了大量的工作. 李群是一种特殊流形, 已被物理学家、化学家广泛使用, 这充分说明在大量的物理、化学数据中蕴涵李群规律, 因此, 用李群方法去分析这些数据的规律已成为一种必然. 所以, 文献[10-12]充分应用计算机科学与人工智能技术方法从机器学习角度出发, 引进李群理论, 给出了一种机器学习新方法, 即李群机器学习的基本概念, 为处理这些复杂的数据提供了新途径.

定义 2(李群机器学习). 一般用 G 表示输入空间, M 表示输出空间. 令 $G \subseteq \mathbb{R}^D$, $M \subseteq \mathbb{R}^d$, $D > d$, 借用李群的定义将 G 对 M 的左作用可用如下映射 φ 表示:

$$\varphi: G \times M \rightarrow M,$$

满足:

- (1) 对于所有的 $x \in M$, $\varphi(e, x) = x$;
- (2) 对于所有的 $g, h \in G$ 和 $x \in M$, $\varphi(g, \varphi(h, x)) = \varphi(gh, x)$.

右作用是满足条件 $\psi(e, x) = x$ 和 $\psi(\psi(x, g), h) = \psi(x, gh)$ 的一个映射 $\psi: M \times G \rightarrow M$.

定义 3(轨道的定义). 若 φ 表示 G 在 M 上作用并且 $x \in M$, 则 x 的轨道定义如下:

$$Orb(x) = \{\varphi_g(x) | g \in G\} \subset M,$$

在有限维情形, $Orb(x)$ 是 M 的浸入子流形. 对于

$x \in M$, φ 在 x 处的稳定(或对称)群由

$$G_x \stackrel{\text{def}}{=} \{g \in G \mid \varphi_g(x) = x\} \subset G$$

给出, 并且 G_x 是一个李子群. $Orb(x)$ 的流形结构由要求双射 $[g] \in G/G_x$ $g \circ x \in Orb(x)$ 为微分同胚而定义, 由此可以判定 G/G_x 是一个光滑流形.

定义 4 (作用的可迁性、有效性和自由性的定义).

(1) 作用具有可迁性, 如果存在唯一的一个轨道, 或等价地, 对于每一个 $x, y \in M$, 有一个 $g \in G$, 使得 $g \circ x = y$;

(2) 作用具有有效性(或者一一对应), 如果 $\varphi_g = id_M$, 蕴涵着 $g = e$, 即 $g \rightarrow \varphi_g$ 是一对一的;

(3) 作用具有自由性, 如果它没有不动点, 即 $\varphi_g(x) = x$ 蕴涵着 $g = e$, 或者等价地, 如果对于每一个 $x \in M$, $\varphi_g = \{e\}$, 并且每一个自由的作用是一对一的.

例 1. 余伴随作用. G 在其李代数 \mathfrak{g} 的对偶 \mathfrak{g}^* 上的余伴随作用定义如下: 设 $Ad_g^*: \mathfrak{g}^* \rightarrow \mathfrak{g}^*$ 是 Ad_g 的对偶, 定义为

$$\langle Ad_g^* \alpha, \xi \rangle = \langle \alpha, Ad_g \xi \rangle,$$

这里 $\alpha \in \mathfrak{g}^*$, $\xi \in \mathfrak{g}$, 则由 $(g, \alpha) \mapsto Ad_{g^{-1}}^* \alpha$ 给出的映射: $\varphi^*: G \times \mathfrak{g}^* \rightarrow \mathfrak{g}^*$ 是 G 在 \mathfrak{g}^* 上的余伴随作用. 相应的 G 在 \mathfrak{g}^* 上的余伴随表示记为 $Ad^*: G \rightarrow GL(\mathfrak{g}^*, \mathfrak{g}^*)$, $Ad_{g^{-1}}^* = (Te(R_g \circ L_{g^{-1}}))^*$.

定义 5 (轨道空间的定义). 等价类的集合 M/G 即称为轨道空间.

定义 6 (学习表达式之间的等价性判定). 设 M 和 N 表示两类学习表达式形成的流形, G 是李群, 并由 $\varphi_g: M \rightarrow M$ 作用在 M 上, 由 $\psi_g: N \rightarrow N$ 作用在 N 上, 构造一个映射 $f: M \rightarrow N$ 关于这些作用是等价的, 如果对于所有的 $g \in G$, 有 $f \circ \varphi_g = \psi_g \circ f$, 如果 M/G 和 N/G 都有正规投影光滑浸没的光滑流形, 则等变映射 $f: M \rightarrow N$ 可以诱导出一个光滑映射 $f_G: M/G \rightarrow N/G$.

3 李群机器学习的公理假设

文献[13-15]给出了李群机器学习的 4 个公理假设, 现归纳如下.

3.1 假设 1: 李群机器学习泛化能力假设公理

该公理包括(1)、(2)、(3)条:

(1) 利用李群、李代数和李子群、李子代数的关系来描述泛化能力, 可定义为: 设 G 是一个李群, 对于 G 的李代数 \mathfrak{g} 的任何一个李子代数 \mathfrak{h} , 存在着唯一

的一个 G 的连通李子群 H , 使得图 1 是可交换的.

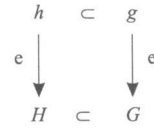


图 1 李群、李代数、李子群、李子代数之间的泛化关系图

(2) 利用学习空间中的内积不变性来描述泛化能力, 可定义为: 设 V 是 n 维线性空间, G 是 $GL(V)$ 的李子群, 如果有 $(g(x), g(y)) = (x, y)$, $\forall x, y \in V$, $g \in G$, 则 V 中内积 (x, y) 称为 G 下不变的.

(3) 利用中心函数来描述泛化能力, 可定义为: 设在紧致李群 G 中取定总体积为 1 的左-右不变的黎曼结构, f 为其上任何一个中心函数, 则 Weyl 积分公式:

$$\int_G f(g) dg = \frac{1}{|w|} \int_{\mathfrak{h}} f(t) |Q(t)|^2 dt,$$

其中 $|w|$ 表示 w 中元素的个数, 当 $t = \exp H$ ($H \in \mathfrak{h}$) 时, $Q(t)$ 可以表示为

$$Q(e^H) = \sum_{\sigma \in w} \text{sign}(\sigma) e^{2xi(\sigma(\delta), H)},$$

其中 δ 是对于范围 C° 的正根之和的一半, 即

$$\delta = \frac{1}{2} \sum_{\alpha \in \Delta^+(G)} \alpha'.$$

3.2 假设 2: 李群机器学习的对偶假设公理

将学习表达式集合构成一个对偶空间来研究学习问题.

例如: 可将李代数 \mathfrak{g} 的对偶 \mathfrak{g}^* 利用泊松括号

$$\{F, G\}(\mu) = \left\langle \mu, \left[\frac{\partial F}{\partial \mu}, \frac{\partial G}{\partial \mu} \right] \right\rangle$$

来研究学习问题. 这种方法已经在物理学中得到广泛的应用.

3.3 假设 3: 李群机器学习的划分独立性假设公理

此公理主要是对学习表达式的分解合理性进行判定.

例如, 任何正交对偶李代数 $(\mathfrak{g}, \alpha, Q)$ 都可以唯一分解成下面的表达式:

$$(\mathfrak{g}, \alpha, Q) = (\mathfrak{g}_0, \alpha_0, Q_0) \oplus (\mathfrak{g}_+, \sigma_+, Q_+) + (\mathfrak{g}_-, \sigma_-, Q_-),$$

其中, $\mathfrak{g}_0 = k_0 \oplus P_0$, $\mathfrak{g}_+ = k_+ \oplus P_+$, $\mathfrak{g}_- = k_- \oplus P_-$, $\alpha_0 = \alpha / \mathfrak{g}_0$, $\sigma_+ = \sigma / \mathfrak{g}_+$, $\sigma_- = \sigma / \mathfrak{g}_-$, $Q_0 = Q / P_0$, $Q_+ = Q / P_+$, $Q_- = Q / P_-$.

3.4 假设 4: 李群机器学习一致性假设公理

主要利用李群的同态关系来研究学习问题. 例如 Q 是李群 G_1 到李群 G_2 上的同态, 则 $Q^{-1}(I_2)$ 是 G_1 的闭正规子群. 又设 π 为 G_1 到 $G_1 Q^{-1}(I_2)$ 的自然

同态, 则有 $G_1/Q^{-1}(I_2)$ 到 G_2 的李群的同态 φ , 使得 $Q=\varphi\pi$.

4 李群机器学习的学习模型

李群机器学习系统中有两种学习模型, 即代数模型和几何模型, 这也就是李群机器学习方法和现有机器学习方法的主要区别之一, 如流形学习、统计学习、支持向量机学习、关系学习、决策树学习、贝叶斯学习等, 根据文献[10-15], 其模型分为代数模型和几何模型.

4.1 李群机器学习的代数模型

在李群机器学习系统中, 将输入数据 R 做变换, 通过 $\theta: R \rightarrow G$ 变换成单参数子群 G , 利用 G 中的右平移变换映射 $R^0: R \times G \rightarrow G; (t, g) \rightarrow g \circ \theta(t)$, 它是流形 G 上的一个可微变换群. 同样可得左平移变换. 由此有(1): {单参数子群} 和 {左不变流形} 是等价的. 再通过左不变流形和左不变向量场之间的等价关系及单参数子群和李子代数之间的关系有 4 类等价关系, 即 $g \cong \{\text{左不变向量场}\} \cong \{\text{左不变流形}\} \cong \{\text{单参数子群}\}$, 即可将李群机器学习的代数模型表示为图 2 所示的形式.



图 2 李群机器学习的代数模型

4.2 李群机器学习的几何模型

主要根据李群的一些几何性质, 如平移不变性、旋转不变性、测地线性质等, 给出李群机器学习的几何模型.

一般来说, 在一个学习系统中, 给定一个观测集, 将观测集映射成一个紧致连通的集合, 这样就可以在观测集中的单位点上取一个自同构 $Ad(G)$ 作用下不变的内积 \langle, \rangle , 对于这样取定的内积, 再取一组标准正交基, 然后用左平移把它们分别扩张成观测集的左不变向量场, 这样就可以唯一地在观测集上的每一个点的切空间取定内积, 使得该内积恰好为该点的切空间的标准正交基, 这样就构成了一个黎曼空间.

由正交向量场组的左不变性可以看出, 所有左平移都是这个黎曼空间的等距变换; 由内积的不变性可知, 所有的右平移也是它的等距变换. 令 $T_e(G)$

为单位点的切空间, 即 G 的李代数 g , $T_a(G)$ 为 G 在任意样本点 a 的切空间, dl_a, dr_a 分别是左平移 l_a 和右平移 r_a 在切空间之间诱导的线性映射, 由此可将李群机器学习的几何模型表示为图 3 所示的形式.

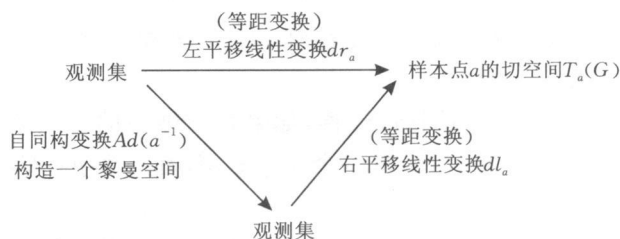


图 3 李群机器学习的几何模型

进一步, 给出李群机器学习的等距变换算法和测地线距离算法.

算法 1. 李群机器学习的等距变换算法.

输入: 样本集 $X = \{x_1, x_2, \dots, x_n \in D\}$

输出: $Y = \{y_1, y_2, \dots, y_n \in d\}, d \leq D$

1. 令输入样本集为 $X = \{x_1, x_2, \dots, x_n \in D\}$, X 满足李群结构.

2. 在 X 的李代数 g 上取定一个在自同构 $Ad(G)$ 的作用下的不变内积:

$$\langle Ad(g)x_1, Ad(g)x_2 \rangle = \langle x_1, x_2 \rangle, \quad x_1, x_2 \in g, g \in G.$$

3. 取定一组标准正交基:

$$\{x_i; 1 \leq i \leq n\}, \quad \langle x_i, x_j \rangle = \delta_{ij}.$$

用左平移把 $\{x_i; 1 \leq i \leq n\}$ 分别张成 X 上左不变向量场, 记为 x_i , 即建立起黎曼空间结构.

4. 在这个黎曼空间上, 对 X 进行左(或右)平移线性变换, 即构成了对样本集的等距映射 $\varphi: X \rightarrow Y, Y = \{y_1, y_2, \dots, y_n \in d\}, d \leq D$.

算法 2. 李群机器学习的测地线距离算法.

输入: 样本点 a

输出: 在 a 的坐标邻域内的所有点与 a 的距离

1. 生成样本点 a 的邻域 U_a .

2. 分析 a 邻域内信息, 在样本集中的单位点上取一个自同构作用下不变的内积 $g_{ij}(a)$, 计算 $g_{ij}(a)$ 的值.

3. 将所得 $g_{ij}(a)$ 的值, 代入

$$D(a, b) = \sqrt{\sum_{i,j=1}^n g_{ij}(a) (a_i - b_i) (a_j - b_j)},$$

计算 U_a 中每一点与 a 之间的距离.

4. 输出所有距离值(通常距离值越小表示与要学习的目标结果越相近).

5 李群机器学习中的 Dynkin 图几何学习算法

文献[12,15-18]主要从几何角度讨论了李群机

器学习中的 Dynkin 图几何学习算法.

5.1 李群机器学习中的 Dynkin 图的描述

在李群机器学习中, 假定 g 是学习空间 G 的复半单李代数, η 为其 Cartan 子代数. g 对 η 的根子空间分解为

$$g = \eta + \sum_{\alpha \in \Delta} g_{\alpha},$$

其中 Δ 为 η 的根系, $|\Delta|$ 为其元素数. η^* 为 η 的对偶空间, 故 $\Delta \subset \eta^*$. 又以 (x, y) 表示 g 的 Killing 型, $\Pi = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 为复半单李代数 g 的素根系. 于是 Π 可以分解为两两正交的子集的和, 而其中每个子集不能再分解:

$$\Pi = \Pi_1 \cup \Pi_2 \cup \dots \cup \Pi_k, (\Pi_i, \Pi_j) = 0, i \neq j \quad (1)$$

自然, Π 的这种分划一定导出 Π 中的一个等价关系, 每个 Π_i 为一个等价类, 这个等价类关系可以用下述方法来建立.

设 $\alpha, \beta \in \Pi$, 如果存在 Π 中序列

$$\alpha = \alpha_{i1}, \alpha_{i2}, \dots, \alpha_{im} = \beta,$$

使得 $\alpha_{ij}, \alpha_{ij+1} \neq 0$ ($j=1, 2, \dots, m-1$), 则称 α 与 β 有关系 \sim , 记为 $\alpha \sim \beta$.

显然, 关系 \sim 有以下性质:

- ① $\alpha \sim \alpha$;
- ② 若 $\alpha \sim \beta$, 则 $\beta \sim \alpha$;
- ③ 若 $\alpha \sim \beta, \beta \sim \gamma$, 则 $\alpha \sim \gamma$.

同一个等价类不能分解为两个相互正交的子集之和; 不同等价类是相互正交的, 即这个等价关系所对应的分划恰为 Π 的分划(式(1)).

定义 7. 若素根系 Π 不能分解为两个非空的相互正交的子集之和, 则称 Π 为不可约素根系, 否则称 Π 是可约的.

由此得到了李群机器学习的问题可解性判别算法 1^[15-17].

定义 8. 设 $\Pi = \{\alpha_1, \alpha_2, \dots, \alpha_l\}$ 是复半单李代数 g 的一个素根系, 以 l 个点分别表示 $\alpha_1, \alpha_2, \dots, \alpha_l$, 将第 i 点与第 j 点用

$$\frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)} \cdot \frac{2(\alpha_i, \alpha_j)}{(\alpha_j, \alpha_j)} = r_{ij}$$

条线段连接起来. 若 $(\alpha_i, \alpha_i) < (\alpha_j, \alpha_j)$, 则画一个从第 j 点指向第 i 点的箭头. 由此得到的图形称为 Π (或 g) 的 Dynkin 图.

5.2 李群机器学习中的 Dynkin 图的相关算法

由上显然, 若 Π 有分解(1), 则 Π 的 Dynkin 图分解为 k 个互不相连的连通分支.

为了研究问题的方便, 我们可以将素根系的性质抽象为空间向量组的性质.

定义 9. 设 E 是一个 n 维的学习空间, $\Pi = \{\alpha_1, \alpha_2, \dots, \alpha_l\}$ 是 E 的一个正向量组, 且当 $i \neq j$ 时, $\frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)} \in \mathbb{Z}^-$. 我们称 $\Pi = \{\alpha_1, \alpha_2, \dots, \alpha_l\}$ 为一个 π -系. 显然有

(1) Π 是线性无关组, 故 $m \leq n$; Π 也是 E 的子空间 $L(\Pi)$ 的 π -系; 当 $m=1$ 时, Π 为 E 的基.

(2) Π 的任何子集 Π_i 也是 π -系.

(3) 若 $\alpha, \beta \in \Pi, \alpha \neq \beta$, 则

$$0 \leq \frac{2(\alpha, \beta)}{(\alpha, \alpha)} \cdot \frac{2(\alpha, \beta)}{(\beta, \beta)} \leq 3.$$

(4) 半单李代数 g 的素根系 Π 是 π -系.

定义 10. 若 π -系 Π 不能分解为两个互相正交的非空子集之和, 则称 Π 是不可约的 π -系.

根据这些结论文献[15-17]给出了李群机器学习的问题可解性判别算法.

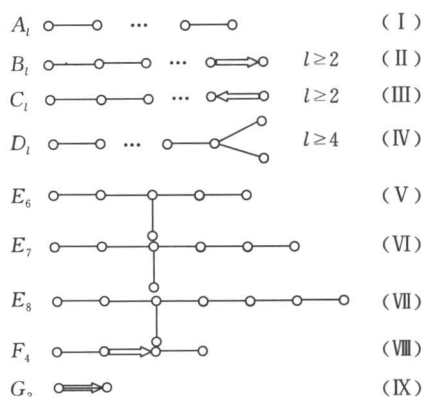
下面引入两个定理.

定理 1. 复半单李代数 g 对应唯一的 Dynkin 图. 证明见文献[15, 19].

定理 2. 两个复半单李代数同构当且仅当它们的 Dynkin 图相同. 证明见文献[15, 19].

文献[15]给出了李群机器学习中的学习空间同构性判别算法.

通过 Dynkin 图的定义和性质, 在同构的定义下, 输入数据向量空间有它对应的 Dynkin 图, 并且只能有 $A_l, B_l, C_l, D_l, E_6, E_7, E_8, F_4$ 与 G_2 . 这 9 种可能^[19] 为



因此, 假设学习空间中有 n 个输入数据, 这些数据用向量组表示为 $\Pi = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$. 我们根据 Dynkin 图对其进行分类, 分类算法如下.

算法 3. 李群机器学习中的 Dynkin 图分类学习算法.

// 初始化: 将输入空间向量组表示为 $\Pi = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$

输入: Π

输出: $\{A_l, B_l, C_l, D_l, E_6, E_7, E_8, F_4, G_2\}$

procedure CLASSIFICATION (II)

if SOLVABLE1(II)=1//学习问题可解

then 以 n 个点分别表示 $\alpha_1, \alpha_2, \dots, \alpha_n$

for $i=1$ to $n-1$; $j=2$ to n do

画线段

$r_{ij} = \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)} \cdot \frac{2(\alpha_i, \alpha_j)}{(\alpha_j, \alpha_j)}$ 连接第 i 点与第 j 点;

if $(\alpha_i, \alpha_j) < (\alpha_j, \alpha_j)$

then 画一个从第 j 点指向第 i 点的箭头;

endif

repeat

endfor

endif

if 有环路

then return;

endif

case // 分类

图(I): A_1 ;

图(II): B_1 ;

图(III): C_1 ;

图(IV): D_1 ;

图(V): E_6 ;

图(VI): E_7 ;

图(VII): E_8 ;

图(VIII): F_4 ;

图(IX): G_2 ;

else return

endcase

end CLASSIFICATION

此算法表明若我们的学习问题是可解的, 那么就可以画出该学习空间对应的 Dynkin 图. 在该算法的基础上, 就可以找到不同类数据对应的李群. 如 A_l 可以对应到特殊酉群 $SU(l+1)$, B_l 可以对应到特殊正交群 $SO(2l+1)$ 等, 再采取对应的问题求解方法, 从而降低了学习问题的求解难度.

6 李群机器学习的分类器设计

一般来说, 线性分类器的构造可以分如下几步来实现.

1. 将样本集映射到 G 这个非空集合上;
2. 根据 G 构造相应的李群结构;
3. 将所得的李群作用于所建立的李群机器学习模型中;
4. 形成相应的分类器;
5. 实例测试;
6. 应用.

关于李群机器学习分类器的研究, 文献[18, 20-

22] 分别给出了量子群、辛群分类器的设计. 现对辛群分类器和量子群分类器做简单介绍.

6.1 李群机器学习中的辛群分类器

6.1.1 基本概念

定义 11. 令 u 和 v 是 $2n$ 维空间 R^{2n} 中的向量, 其中

$$u = (x_1, \dots, x_n, \xi_1, \dots, \xi_n)^T,$$

$$v = (y_1, \dots, y_n, \eta_1, \dots, \eta_n)^T,$$

则 R^{2n} 中的辛内积定义为

$$\omega(u, v) = \sum_{i=1}^n (x_i \eta_i - y_i \xi_i) = u^T J v^T,$$

其中 $J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$, 具有性质 $J^{-1} = J^T = -J$, I_n

为 n 阶单位阵, 辛内积的基本性质如下:

(1) 双线性. $\omega(\lambda_1 u_1 + \lambda_2 u_2, v) = \lambda_1 \omega(u_1, v) + \lambda_2 \omega(u_2, v)$;

(2) 反对称. $\omega(u, v) = -\omega(v, u)$;

(3) 非退化. 对于每一个 $v \in R^{2n}$, 如果均有 $\omega(u, v) = 0$, 则 $u = 0$.

定义 12. 具有上述条件的辛内积 ω 的实向量空间 (R^{2n}, ω) , 称为辛空间.

定义 13. 如果辛空间 (R^{2n}, ω) 中的线性变换 S 满足 $S: R^{2n} \rightarrow R^{2n}$, 对于一切 $u, v \in R^{2n}$ 均有

$$\omega(Su, Sv) = \omega(u, v),$$

则线性变换 S 为辛变换或正则变换, 辛变换保持两向量的辛内积不变.

定义 14. 设 (V, ω) 是一辛空间, 所谓 (V, ω) 的一个自同构, 是指 (V, ω) 到其自身的一个同构, 所以 (V, ω) 是 V 的线性变换群 $GL(V)$ 的一个元素, 若把它记为 s , 则它满足

$$\omega(sx, sy) = \omega(x, y), \quad \forall x, y \in V,$$

易知 (V, ω) 的自同构全体辛变换群构成群 $GL(V, R)$ 的一个子群, 我们把它记为 $Sp(V, \omega)$. 特别地, 标准辛空间 (k^{2n}, ω) 的自同构群记为 $Sp(2n, k)$, 若 $k = R$, 则把 $Sp(2n, k)$ 简记为 $Sp(2n)$, 并称它为 $2n$ 维辛群^[23-24].

由此有以下定理.

定理 3. 辛空间 (R^{2n}, ω) 一定是偶数维的. 证明见文献[22].

6.1.2 辛群分类器的设计方法

在辛群分类器的设计中, 首先应该将分类样本表示成辛群形式, 然后利用辛群的张量表示方法, 将样本构造成对应的辛矩阵, 对辛矩阵进行化简求得对应的训练特征, 以此作为学习过程训练样例. 最后

对待检测样本按同样方式进行处理, 利用辛群分类算法得到样例特征, 如果两类特征匹配, 则将结果输出. 接下来给出辛群分类器算法^[42].

算法 4. 辛群分类器算法.

输入: 样例集 D

输出: 某类特殊的辛矩阵 $\{M(1), M(2), M(3), M(4), M(5)\}$

1. 将样例数据集 D 映射到 $Sp(2n)$ 中, 取 $Sp(2n)$ 上的 n 维行向量组成集合 S_q^a , $S_q^a = \{(x_1, x_2, \dots, x_n) | x_i \in S_q, i = 1, 2, \dots, n\}$.

// $n = 2v$, v 是正整数, 向量 (x_1, x_2, \dots, x_n) 表示任意一个样本

2. 选取 $Q_i \in Sp(2n)$. 对于 Q_i 将它转换成对应的辛矩阵, 并求出该辛矩阵的奇异值.

3. 将 Q_i 作用于样本数据集 $((x_1, x_2, \dots, x_n), Q_i)$ $(x_1, x_2, \dots, x_n)Q_i$ 作为学习过程中的训练样例.

4. 在训练阶段, 对训练样本 X_{ji} 做奇异值分解. 取前 k 个最大奇异值对应的样本 $A_{ji}(m) = u_{ji}^m (v_{ji}^m)^H$ ($m = 1, 2, \dots, k$), 由训练样本 X_{ji} 得到的样本构成了第 j 类的一个判别函数 $g_{ji}(X)$,

$$g_{ji}(X) = \sum_{m=1}^k \alpha_{ji}^2(m) = \sum_{m=1}^k |\langle A_{ji}(m), x \rangle|^2,$$

$j = 1, 2, \dots, a$ $i = 1, 2, \dots, n$, 其中 $\alpha_{ji}(m) = \langle A_{ji}(m), X \rangle$ 为待识别样本图像 X 到第 j 类的第 i 个训练样本 X_{ji} 的第 m 个基图像 $A_{ji}(m)$ 的投影值.

5. 在样例集 D 中删除已获得的目标学习结果的样例, 获得新的样例集 D' .

6. 将待识别样本 X 输入每个判别函数 $g_{ji}(X)$, 求这个判别函数的输出, 将 X 归入最大输出的 $g_{ji}(X)$ 所对应的类

别中, 若每类有多个训练样本, 首先对每类计算最大长度 $h_j = \max_i (g_{ji}(X))$ ($j = 1, 2, \dots, a$, $i = 1, 2, \dots, n$), 这样就可以将 X 归入到第 c 个类别中.

7. 对每个目标函数 $t(x)$, 输出集合 F .

注: 设 $M = \begin{bmatrix} A & B \\ A_2 & B_2 \end{bmatrix} \in Sp(2n)$, 有

$$(1) M \in Sp(2n) \cap O_{2n} = U_n \Rightarrow M = \begin{bmatrix} A & J^{-1}A \\ A^{-1}JA & O_n \end{bmatrix}, A^{-1}A = I_n \quad M(1)$$

$$(2) M = \begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix} \in Sp(2n) \Rightarrow M = \begin{bmatrix} A_1 & 0 \\ 0 & A_1^{-1} \end{bmatrix}, A_1 \in GL(n, R) \quad M(2)$$

$$(3) M = \begin{bmatrix} I & 0 \\ 0 & B_2 \end{bmatrix} \in Sp(2n) \Rightarrow M = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, S^{-1} = S \quad M(3)$$

$$(4) M = \begin{bmatrix} A_1 & 0 \\ A_2 & I \end{bmatrix} \in Sp(2n) \Rightarrow M = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, S^{-1} = S \quad M(4)$$

$$(5) M = J_a = \begin{bmatrix} I\hat{\alpha} & I\alpha \\ -I\hat{\alpha} & I\alpha \end{bmatrix}, \alpha \subset v \in Sp(2n) \quad M(5)$$

6.1.3 实例分析

我们从 ORL 通用人脸图像测试库中, 就人脸的 7 种姿态, 即正面、俯、仰、左平面内侧、左平面外侧、右平面内侧和右平面外侧进行处理, 并和用 SVM 算法^[25] 进行比较, 其结果为表 1.

表 1 SVM 分类平均识别率和辛群分类平均识别率比较表

算法	平均识别率				
	正面图像/%	俯视图像/%	仰视图像/%	左平面内侧图像/%	右平面内侧图像/%
SVM 分类算法	95.4	93.8	90.2	93.4	93.4
辛群分类算法	95.2	92.6	93.1	94.6	94.6

实验表明, 当人脸图像经过特征提取后, 在同一个人脸图像较多时, 辛群分类算法在分类正确率上有一定优势, 在样例较多时, 具有较好的分类效果.

6.2 李群机器学习中的量子群分类器

量子群是经典李群、李代数的基本对称概念的推广, 与非交换几何、量子对称性等密切相关, 针对机器学习系统中面对的非交换性和非对称性问题, 文献[21] 提出了量子群分类器的构造方法.

6.2.1 量子群分类器的构造

构造量子群分类器, 只要利用量子群的 Jacobi 条件分析观测数据的量子群和量子代数的性质, 并将量子群看成量子超平面的线性系统, 通过非交换空间上的降维、线性化等形式对观测数据进行处理,

将观测数据的非线性结构约简成线性结构, 然后根据量子群上线性变换学习算法对其进行分类.

6.2.2 量子群的对称线性变换学习算法

在二维量子超平面上, 假设学习空间中的学习问题用“坐标” (x, y) 表示, 其中 x 表示学习系统中的实例空间中的实例, y 表示样本空间中的样例, 且满足下述交换关系:

$$xy = qyx, q \in C \quad (2)$$

其中, q 表示调整系数, C 是调整系数的集合. 此外, (x, y) 还存在相关的协变微分计算, 即存在“形式” (ξ, η) , 它们满足

$$\xi\eta = -q^{-1}\eta\xi, \xi^2 = 0, \eta^2 = 0 \quad (3)$$

当 $q = 1$ 时, x 与 y 对易, ξ 与 η 反对易, (ξ, η) 可以解

释为对 (x, y) 的微分,也即对学习空间中的实例进行降阶处理.令

$$\phi=\begin{bmatrix} x \\ y \end{bmatrix}, \quad \psi=d\phi=\begin{bmatrix} \xi \\ \eta \end{bmatrix} \tag{4}$$

量子超平面上“坐标” (x, y) 与“形式” (ξ, η) 同时作线性变换:

$$\begin{bmatrix} x' \\ y' \end{bmatrix}=\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}, \quad \begin{bmatrix} \xi' \\ \eta' \end{bmatrix}=\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}\begin{bmatrix} \xi \\ \eta \end{bmatrix} \tag{5}$$

这里设 $(\alpha, \beta, \gamma, \delta)$ 与 (x, y, ξ, η) 相互可交换,如要求变换后 (x', y') 与 (ξ', η') 仍满足与 (x, y) 及 (ξ, η) 间相同的关系,则 $(\alpha, \beta, \gamma, \delta)$ 间必有如下关系成立:

$$\begin{aligned} \alpha\beta &= q\beta\alpha, \quad \alpha\gamma = q\gamma\alpha, \quad \beta\gamma = q\gamma\beta, \quad \gamma\delta = q\delta\gamma, \\ \beta\delta &= q\delta\beta, \quad \alpha\delta = (q - q^{-1})\beta\gamma \end{aligned} \tag{6}$$

令代数

$$A=C\langle \alpha, \beta, \gamma, \delta \rangle \sim$$

为由 $(\alpha, \beta, \gamma, \delta)$ 生成的可结合代数,且满足关系(6),即

$$T=(T_i^j)=\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \in GL_q(2),$$

则上述变换便构成量子超面的对称线性变换,得到量子群 $GL_q(2)$,其中心为

$$c=det_q T=\alpha\delta - q\beta\gamma = \delta\alpha - q^{-1}\beta\gamma.$$

如要求 $det_q T=1$,则称为量子群 $SL_q(2)$.

现将量子群上的对称线性变换学习算法描述如下.

算法 5. 量子群上的对称线性变换学习算法.

输入: 实例集 $X=\{x_1, x_2, \cdots, x_n\}$,

样本集 $Y=\{y_1, y_2, \cdots, y_n\}$

输出: 量子群的对称线性变换

1. 对 X 和 Y 进行降维处理后,得到 ξ 与 η ,且它们必须满足关系(3);
2. 对 X, Y 和 ξ, η 分别同时作对称线性变换 Z ,使得变换后的 X, Y 和 ξ, η 仍满足关系(2)和(3);
3. 检验对称线性变换 Z 是否满足关系(6),若为否,则调整系数 q ,直到找到合格的变换 Z ,则 Z 即为所求的对称线性变换.
4. 返回线性变换 Z .

6.2.3 量子群分类器在 DNA 序列分类中的应用

在介绍具体的 DNA 序列分类实验之前,先对实验数据作一个简单的说明^[21, 43-47].

我们选择真核生物中拟南芥(*Arabidopsis thaliana*)的完全序列作为实验数据.数据来源于 NCBI(National Center for Biotechnology Information),其网址为 <http://www.ncbi.nlm.nih.gov>.该全序列共有 5 条染色体,其中每一条染色体都是

一个 FAST 格式的文件,由于篇幅原因在此省略 FASTA 格式文件的介绍,只简单说明一下各染色体的索引 ID、长度及文件大小,见表 2.

表 2 染色体文件说明

染色体	索引 ID	长度/ bp	大小/ M
1	NC_003070	30080809	43
2	NC_003071	19643621	27.7
3	NC_003074	23465812	33.5
4	NC_003075	17549528	24.9
5	NC_003076	26689408	38

在信息提取过程中,采用下面的结构体组织存储从每个染色体序列中读出的信息:

```
typedef struct
{
    char * seqname;           // 基因序列名称//
    char * wholeseq;          // 完整的基因序列//
    char * info;              // 基因附加信息//
    long index;               // 本基因序列在基因库中的位置//
    unsigned len;             // 基因序列的长度//
    int * count;              // 各个观测值在基因序列中的个数//
}SeqInfo;
```

在提取数据时,我们的主要目的是将每个染色体序列中的编码区和非编码区分别提取出来.采用的方法是:

- (1)对每条染色体,读取并保存 gene 后的起始及结尾碱基号和 CDS 后各段的始末碱基号;
- (2)在最后的序列中找到 CDS 中标志的位置,将这些段连接起来作为第 i 个 SeqInfo 结构体存入编码区文件 coding.txt 中;
- (3)将剩下的基因段连接起来,作为第 i 个 SeqInfo 结构体存入非编码区文件 noncoding.txt 中.
- (4)重复上述步骤,直到最后一个染色体.

对于 5 条染色体,我们分别得到 5 个存放编码区的文件和 5 个存放非编码区的文件.各个文件的数据数目见表 3.

表 3 各染色体编码区与非编码区的数据数目

染色体	编码区数据大小	非编码区数据文件
1	6761	6572
2	4216	4134
3	5252	5027
4	3958	3482
5	5987	5160

由于编码区和非编码区有着不同的功能,因此

会有一些序列在编码区中频繁出现,而在非编码区出现得很少,这些词就可以被称为编码区的特征序列,相反则视为非编码区的特征序列.由于每个 DNA 序列都是由 A、T、G、C 4 个字符随即组成的字符串(如 aggcacggaaaacgggaatt),数字化 DNA 序列的方法很多,不同的方法对于分类的影响很大.

在进行 DNA 序列处理之前一般是先寻找一种 DNA 序列的数学表示,再借助于其它工具对其进行分析和研究.现在国际上有许多 DNA 序列的数学表达方法,如随机步模型;DNA 步、基于混沌理论 CGR 方法以及利用几何方法的 Z 曲线表示法等等.本文我们采用基于序列的熵值的方法,由于该序列集共有 5 条染色体,可以分别在每条染色体中提取特征序列,然后对每个基因中的编码和非编码区应用我们所设计的量子群分类器进行分类.

首先,将一串 DNA 序列看成一个信息流,考虑其单位序列所含信息量(即熵)的多少.从直观上理解,我们认为,重复得越多,信息量越少.

设序列 $L=(a_1,a_2,\cdots,a_n)$,前 m 个字符所带的信息量为 $f_m(L)$,记为

$$g_m(L)=f_m(L)-f_{m-1}(L) \tag{7}$$

其中, $g_m(L)$ 为加上第 m 个字母之后所增加的信息量.然后,由式(7),得 $f_n(L)=\sum_{i=1}^n g_i(L)$,则 $f_n(L)$ 为整个序列所带得信息量, $\frac{f_n(L)}{|L|}$ 即为单位长度所带得信息量.现在问题就归结为如何找出一个合适的 $g_m(L)$.

对此,我们认为 $g_m(L)$ 应具备以下性质.

性质 1. $g_m(L)>0$,即任意加上一个字符,它或多或少带有一定的信息量;

性质 2. 第 m 个字符(或者是以它结尾的较短序列)与前面的序列(信息流)重复得越多, $g_m(L)$ 的值就越小;

性质 3. 第 m 个字符(或者是以它结尾的较短序列)如果和与它靠得越近的重复, $g_m(L)$ 的值就越小,反之,则越大;

性质 4. $f_0(L)=0$.

因此,我们构造如下的函数:

$$g_m(L)=\frac{b}{b+t_1\sigma_1+t_2\sigma_2+\cdots+t_p\sigma_p},$$

其中, b 为防止分母为零而设置的一个小正数, $\sigma_i=$

$$\sum_{i=1}^n a^i \delta_i, \text{ 且}$$

$$\delta_i=\begin{cases} 1, & \text{以第 } m-t \text{ 个字符结尾的 } i \text{ 字串且以} \\ & \text{第 } t \text{ 个字符结尾的 } i \text{ 字串完全相同} \\ 0, & \text{否则} \end{cases},$$

其中 a 为一个小于 1 的数, p 取为某一固定的正数.经过反复上机搜索,我们取 $p=6$,即只检查长度为 1~6 的字串即可.因为,字串长度太大的重复非常少见.另外,取 $a=0.392, b=0.1, c=3$.

借助于 MATLAB,我们提取出由 A、T、G、C 4 个碱基组成的字符串的特征序列,各条染色体的特征序列个数分别为 100、76、86、76、93.

在量子超平面上,设有 n 个特征序列 $\{x_i\}$,它们是长度不同的特征序列的集合,对一条长度为 L 的 DNA 序列,通过下列计算我们可以得到这条 DNA 序列的数学表示:

$$V_{ij}=\lambda \cdot \frac{T_i}{L-l_i+1} \cdot \frac{l_i}{l_j}, \quad i, j=1, 2, \cdots, n,$$

其中 V_{ij} 为群 V 的第 i 行 j 列的元素,且 V 中元素的按行优先原则按照长度由小到大排列,长度相同的按照字符的顺序排列, T_i 为第 i 个特征序列在这条 DNA 序列中出现的总次数, l_i, l_j 分别为第 i 个和第 j 个特征序列的长度, λ 是一个放大系数,为了避免分量值过小引起计算溢出.

构造出群 V 之后,判断出 V 满足量子群的定义,是一个量子群,然后将量子群的对称线性变换学习算法 5 作用于量子群 V 上,便可得到我们所需的经典量子群的结构,再利用量子群分类器的分类算法便可进行分类.

实验中,我们对每条染色体,将数据平均分为 5 段,每次随机取 2 段作为训练集,得出判断函数,剩下的 3 段作为测试集.我们将对 5 条染色体采用量子群分类器分类的精确度与 SVM 算法的分类精确度分别进行比较,如表 4 所示.

表 4 量子群分类器与 SVM 分类效果比较

染色体编号	SVM 分类精度(100%)	量子群分类器分类精度(100%)
1	97.34	97.58
2	92.29	93.44
3	94.76	95.25
4	92.21	92.57
5	97.31	97.93

在实验的特征提取过程中,我们考虑的是每条 DNA 所蕴含的信息量,只考虑了某一方面的特征,所以,总有一些不尽如人意的地方.应该将单个字母出现的频率、周期性等因素综合起来考虑,同时考虑序列中元素的局部性质和序列的全局性质,使序列各方面的特征都能得到体现,使分类更加科学.

从实验结果来看,在第 2 和第 4 条染色体上,本文的分类方法与 SVM 的方法精确度都较低,这是因为第 2 和第 4 条染色体的上提取的特征序列相对较少,但我们的方法相对 SVM 方法精度却有很明显的提高,而且在第 1、3、5 染色体上,两者的精确度比较相近.实验证明本文所提出的量子群分类器,即使在特征序列相对较少的情况下也能很好地进行分类,且分类精度也相对较高.

7 李群机器学习子空间轨道生成学习算法

关于这方面的工作文献[26-30]分别提出了:

(1) LML 中偏序集及格的基本概念;

(2) LML 子空间轨道生成格的广度、深度优先及启发式学习算法;

(3) 一般线性群 $GL_n(F_q)$ 作用下学习子空间轨道生成学习算法及辛群作用下子空间轨道生成学习算法等,并将这些算法进行了初步应用,取得了满意的结果.

8 结论与展望

综上所述,本文对目前取得的李群机器学习的相关成果进行总结.理论上取得了一些有意义的成果;应用上已分别在计算机图形学、数据分类、药物分子设计、晶体分类、人脸识别等方面^[42-48]得到了初步应用.为了让更多的研究者参与到该领域来,下面列出一些研究方向,供有兴趣的研究者参考:

(1) 李群学习表达式映射机制问题,如学习表达式的表示、学习表达式之间的关系及表达式关系之间的关系等;

(2) 李群深层结构学习机制问题:如泊松结构学习、辛结构学习等;

(3) 更多的现代数学方法应用在机器学习领域,如群范畴方法、层范畴方法等;

(4) 物理学方法在机器学习中的应用等.

参 考 文 献

- [1] Seung H S, Lee D D. The manifold ways perception. Science, 2000, 290: 2268-2269
- [2] Roweis S, Saul L. Nonlinear dimensionality reduction by locally linear embedding. Science, 2000, 290: 2323-2326

- [3] Tenenbaum J, Silva V D, Langford J. A global geometric framework for nonlinear dimensionality reduction. Science, 2000, 290: 2319-2323
- [4] Donoho D L, Grimes C. Hessian eigenmaps: New locally linear embedding techniques for high-dimensional data. Proceedings of the National Academy of Sciences, 2003, 100 (10): 5591-5596
- [5] Zhang C S, Wang J, Zhao N Y, Zhang D. Reconstruction and analysis of multi-pose face image based on nonlinear dimensionality reduction. Pattern Recognition, 2004, 37(1): 325-336
- [6] Belkin M, Niyogi P. Laplacian eigenmaps for dimensionality reduction and data representation. Neural Computation, 2003, 15(6): 1373-1396
- [7] Zhang Z Y, Zhang H Y. Principal manifolds and nonlinear dimensionality reduction via tangent space alignment. SIAM Journal of Scientific Computing, 2005, 26(1): 313-338
- [8] Zhan De-Chuan, Zhou Zhi-Hua. Ensemble-based manifold learning for visualization. Journal of Computer Research and Development, 2005, 42(9): 1533-1537(in Chinese)
(詹德川, 周志华. 基于集成的流形学习可视化. 计算机研究与发展, 2005, 42(9): 1533-1537)
- [9] Yang Jian, Li Fu-Xin, Wang Jue. A better scaled local tangent space alignment algorithm. Journal of Software, 2006, 16(9): 1584-1590(in Chinese)
(杨剑, 李伏欣, 王珏. 一种改进的局部切空间排列算法. 软件学报, 2006, 16(9): 1584-1590)
- [10] Xu Huan, Li Fan-Zhang. Study on Lie group machine learning. Journal of Computational Information Systems, 2005, 1(4): 843-849
- [11] Li Fan-Zhang, Kang Yu. The study of machine learning theory frame based on Lie group. Journal of Yunnan Nationalities University, 2004, 13(4): 251-255(in Chinese)
(李凡长, 康宇. 基于 Lie 群的机器学习理论框架. 云南民族大学学报, 2004, 13(4): 251-255)
- [12] Li Fan-Zhang, Xu Huan. The theory framework of Lie group machine learning (LML). Computer Technology and Application, 2007, 1(3): 62-80
- [13] Xu Huan, Li Fan-Zhang. Lie group machine Learning's axioms hypothesizes//Proceedings of the 2006 IEEE International Conference on Granular Computing, 2006: 401-404
- [14] Li Fan-Zhang. Research on the Lie group machine learning. School of Computer Science and Technology, Suzhou University, Technical Report, 2004
- [15] Xu Huan. Research and application on Lie-group Machine Learning models[M. S. dissertation]. Suzhou University, Suzhou, 2007(in Chinese)
(许欢. 李群机器学习模型及其应用研究[硕士学位论文]. 苏州大学, 苏州, 2007)
- [16] Xu Huan, Li Fan-Zhang. Geometry algorithms of dynkin diagrams in Lie group machine learning. Journal of Nanchang Institute of Technology, 2006, 25(2): 75-78

- [17] Xu Huan, Li Fan-Zhang. Algorithms of dynkin diagrams in Lie group machine learning. *Journal of Communication and Computer*, 2006, 4(3): 13-17
- [18] Xu Huan, Li Fanzhang. The design of $SU(n)$ classifier of Lie group machine learning (LML). *Journal of Computational Information Systems*, 2005, 1(4): 835-841
- [19] Xiang Wu-Yi, Hou Zi-Xin, Meng Dao-Ji. *Lectures on Lie Group*. Beijing: Peking University Press, 1992(in Chinese)
(项武义, 侯自新, 孟道骥. 李群讲义. 北京: 北京大学出版社, 1992)
- [20] Li Fanzhang, Xu Huan. $SO(3)$ Classifier of Lie group machine learning. *Journal of Communication and Computer*, 2006, 4(3): 76-80
- [21] He Shu-Ping. Research on quantum group classifier in Lie-group Machine Learning[M.S. dissertation]. Suzhou University, Suzhou, 2008(in Chinese)
(何书萍. 李群机器学习中的量子群分类器研究[硕士学位论文]. 苏州大学, 苏州, 2008)
- [22] Fu Hui-Xin. Research on symplectic group classifier in Lie group machine learning[M.S. dissertation]. Suzhou University, Suzhou, 2008(in Chinese)
(付会欣. 李群机器学习中的辛群分类器研究[硕士学位论文]. 苏州大学, 苏州, 2008)
- [23] Jiang Le-Le, Wu Guang-Liang. Study on the scattering of two-dimensional electromagnetic field based on symplectic integrator method. *Acta Electronica Sinica*, 2004, 32(12): 1967-1970(in Chinese)
(蒋乐乐, 吴光良. 基于辛算法的二维电磁场散射问题的研究. 电子学报, 2004, 32(12): 1967-1970)
- [24] Hou Zi-Xin. New examples of almost paracontact manifolds. *Advances in Mathematics*, 1993, 22(4): 354-357(in Chinese)
(侯自新. 具有殆仿切触结构流形的新例子. 数学进展, 1993, 22(4): 354-357)
- [25] Shen Xin-Yu, Yu Hong-Li, Guan Teng-Fei. Image classification based on transductive support vector machines. *Journal of Computer Applications*, 1993, 22(4): 354-357(in Chinese)
(沈新宇, 许宏丽, 官腾飞. 基于直推式支持向量机的图像分类算法. 计算机应用, 2007, 26(6): 1463-1467)
- [26] Chen Feng, Li Fan-Zhang. Orbits generated theory of learning subspace and its algorithm in Lie-group Machine Learning (LML). *Journal of Suzhou University*, 2007, 23(1): 61-66(in Chinese)
(陈凤, 李凡长. 李群机器学习(LML)的学习子空间轨道生成理论及算法初探. 苏州大学学报, 2007, 23(1): 61-66)
- [27] Chen Feng, Li Fan-Zhang. Orbits generated lattice algorithm of learning subspace by $GL_n(F_q)$ action in Lie-group Machine Learning (LML)// *Proceedings of the 1st Annual Conference of Jiangsu Young Scientist*, 2006: 36-46(in Chinese)
(陈凤, 李凡长. 李群机器学习(LML)中 $GL_n(F_q)$ 群作用下的学习子空间轨道生成格算法研究// 首届江苏省青年科学家年会会议论文集, 2006: 36-46)
- [28] Chen Feng, Li Fan-Zhang. Orbits generated lattice algorithm of learning subspace in Lie-group Machine Learning (LML). *Computer Engineering and Applications*, 2007, 43(15): 184-187(in Chinese)
(陈凤, 李凡长. 李群机器学习(LML)的学习子空间轨道生成格算法研究. 计算机工程与应用, 2007, 43(15): 184-187)
- [29] Chen Feng. Orbits generated lattice algorithm of learning subspace in Lie-group Machine Learning (LML) and application research [M. S. dissertation]. Suzhou University, Suzhou, 2007(in Chinese)
(陈凤. 李群机器学习(LML)子空间轨道生成算法及其应用研究[硕士学位论文]. 苏州大学, 苏州, 2007)
- [30] Li Fan-Zhang, He Shu-Ping, Lin T Y. Lie group machine learning// *Proceedings of the 2nd Annual Conference of Jiangsu Young Scientist*. Nanjing, 2008: 44
- [31] Rao Rajesh P H, Ruderman Daniel L. Learning Lie groups for invariant visual perception. *Advances in Neural Info Processing Systems*, 1999, (11): 810-816
- [32] Govindu Venu Madhav. Lie-algebraic averaging for globally consistent motion estimation// *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004, 1: 1684-1691
- [33] Squire David McG. Invariance signatures: Characterizing contours by their departures from invariance. *Computer Vision and Image Understanding*, 2000, 77: 284-316
- [34] Arnold D Gregory, Sturte Kirk, Veltin Vince, Nadhakumar N. Dominant-subspace invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(7): 649-662
- [35] Tsao Tien-Ren, Shyu Haw-Jye, Libert John M, Chen Victor C. A Lie group approach to a neural system for three-dimensional interpretation of visual motion. *IEEE Transactions on Neural Networks*, 1991, 2(1): 149-155
- [36] Fiori Simone. Unsupervised neural learning on Lie group. *International Journal of Neural System*, 2002, 12(13): 1-28
- [37] Nishimani Yasunori, Akaho Shotaro. Learning algorithms utilizing quasi-geodesic flows on the stiefel manifold. *Neural Computing*, 2005, 67: 106-135
- [38] Mansouri Abdol-Reza, Mukherjee Dipti Prasad, Action Scott T. Constraining active contour evolution via Lie groups of transformation. *IEEE Transactions on Image Processing*, 2004, 13(6): 853-863
- [39] Kantor George A, Krishnaprasad P S. An application of Lie groups in distributed control networks. Preprint Submitted to Elsevier Preprint, 2000
- [40] Lebanon Guy. Metric learning for text documents. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(4): 1-12
- [41] Kondor Risi. Group theoretical methods in machine learning [Ph. D. dissertation]. Columbia University, New York, 2008

- [42] Fu Hui-Xin, Li Fan-Zhang. Research of the symplectic group classifier based on Lie group machine learning//Proceedings of the 2007 IEEE the 4th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD' 07). 2007: 649-655
- [43] He Shu-Ping, Li Fan-Zhang. A symmetrical linear transform learning algorithm on quantum group//Proceedings of the 2007 IEEE International Conference on Granular Computing, 2007: 724-727
- [44] He Shu-Ping, Li Fan-Zhang. A molecular docking drug design algorithm based on quantum group. Journal of Nanjing University, 2008, 44(5): 512-519(in Chinese)
(何书萍, 李凡长. 一个基于量子群的分子对接药物设计算法. 南京大学学报, 2008, 44(5): 512-519)
- [45] Zhang Jiong, Li Fan-Zhang. A fiber bundle model based on manifold learning. Journal of Nanjing University, 2008, 44(5): 477-485(in Chinese)
(张炯, 李凡长. 基于流形学习的纤维丛学习模型研究. 南京大学学报, 2008, 44(5): 477-485)
- [46] Guan Wen-Wen, Li Fan-Zhang. Drug molecular design using Lie group machine learning (LML)//Proceedings of the 2008 International Conference on Advanced Intelligent. Beijing, China, 2008: 411-414
- [47] Xian Min, Li Fan-Zhang. A homology coboundary learning algorithm for data reduction//Proceedings of the 2008 International Conference on Advanced Intelligent. Beijing, China, 2008: 97-102
- [48] Fiori Simone. Leap-frog-type learning algorithms over the Lie group of unitary matrices. Neurocomputing, 2008, 71(10-12): 2224-2244
- [49] Fiori Simone. Lie-group-type neural system learning by manifold retractions. Neural Network, 2008, 21(10): 1524-1529
- [50] Xu Han-Xiang, Li Fan-Zhang. Semi-supervised learning algorithm based on Lie group//Proceedings of the GIC S2009. 2009, 3: 573-577
- [51] Fiori Simone. Learning-machine-committee averages over the unitary group of matrices//Proceedings of the ISCAS2009. 2009: 2777-2781
- [52] Lin Tong, Zha Hong-Bin. Riemannian manifold learning. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(5): 796-809
- [53] Li Fan-Zhang, Qian Xu-Pei, Xie Lin, He Shu-Ping. Machine Learning and its Applications. Hefei: University of Science and Technology of China Press, 2009(in Chinese)
(李凡长, 钱旭培, 谢琳, 何书萍. 机器学习理论及应用. 合肥: 中国科技大学出版社, 2009)
- [54] Tuzel O, Porikli F, Meer P. Pedestrian detection via classification on riemannian manifolds. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(10): 1713-1727



LI Fan-Zhang professor, Ph. D. supervisor. His main research interests include Lie group machine learning, dynamic fuzzy logic.

HE Shu-Ping master, teaching assistant. Her main research interests include Lie group machine learning.

QIAN Xu-Pei master, lecturer. His main research interests include Lie group machine learning.

Background

As a new approach to learning in machine learning field, Lie group machine learning, on the one hand, keeps the merits of manifolds; On the other hand, it borrows the idea of Lie group. And so a learning paradigm with innovative feature was formed. Since reported in 1994, Lie group machine learning has draw widespread attention at home and abroad, such as Finland, Canada, Ireland, Italy, USA. Compared with manifold leaning method, it has an obvious advantage. As can be seen from the concept of Lie Group, it includes the contents of differentiable manifold and group; differentiable

manifold contains topological manifold and differential structure. In addition, seen from cognitive process, when identifying any object in the objective world, human brain first look to a stable point in symbolizing problems, second analyzes the image structure in turn. In Lie group structure, minimum generator is this stable point for the cognitive models. As long as the minimum generator is found, one can analyses the image by Lie group method. So Lie group learning method not only is up to the cognitive and leaning rule, but also meets the conditions of solving realistic problem by computer.