高维希尔伯特度量中的分类器

Aditya Acharya 马里兰大学帕克分校计算机科学系，美国

Auguste H. Gezalyan 马里兰大学帕克分校计算机科学系，美国

David M. Mount
马里兰大学帕克分校计算机科学系，美国

arXiv:2601.13410v1 [cs.CG] 19 Jan 2026

摘要

在高维空间中对点进行分类，是机器学习中的一个基础几何问题。本文研究在 d 维希尔伯特多边形度量下对点进行分类。希尔伯特度量是凯莱-克莱因双曲距离向任意凸体的推广，在机器学习和凸几何中有广泛应用。我们首先针对大间隔支持向量机问题，提出一种在该度量下高效的基于线性规划的算法。该算法的运行时间关于点的数量、边界面数量和维数均为多项式。与以往的研究相比，这是一个重大改进；以往的研究要么没有提供运行时间的理论保证，要么运行时间呈指数级。我们还研究了密切相关的芬克度量，并针对软间隔支持向量机问题以及希尔伯特度量下基于最近邻的分类，提出了高效算法。

2012年美国计算机协会学科分类 计算理论 → 计算几何

Ke关键词和短语 支持向量机、希尔伯特几何、分类、机器学习 ng

数字对象唯一标识符 10.4230/LIPIcs...

1 引言

在机器学习中，二元分类 是一项监督学习任务，算法通过学习将输入数据（特征点）映射到两个互斥类别或标签之一，通常分别记为正类 和负类。其主要目标是训练一个模型，使其能够准确预测一个新的、未见过的数据点属于两个预定义类别中的哪一个。二元分类是机器学习中最基础的几何问题之一 [6, 9, 28, 29]，因为它与广泛的经验决策问题相关。

支持向量机（SVM）由 Vapnik 和 Chervonenkis [31]提出，是一种二元分类算法 [8]。给定两组训练点P^+ 和 P^-，它们表示从两个不同空间分布中采样得到的特征向量。支持向量机通过计算一个超平面来分隔这两组训练集，从而求解这一分类问题；从每组点到分离超平面的最小距离称为间隔，并且该距离应尽可能大。

支持向量机广泛应用于计算机视觉 [19, 30], 自然语言处理 [11, 15], 和计算生物学 [3, 13] 等众多领域。它们在欧几里得几何和双曲几何 [7], 中已得到充分研究，但在其他非欧几里得几何中的分类研究相对较少。本文从希尔伯特几何及其密切相关的芬克几何的角度研究这一问题。

希尔伯特度量（将在第 2 节中定义）是将双曲几何的凯莱–克莱因模型推广到任意凸集。它具有若干理想性质，例如测地线是直线，且在射影变换下保持不变 [26]。希尔伯特度量有广泛的应用，包括：量子信息论 [27], 实分析 [5], 以及机器学习 [17, 24, 25]。请注意

© Aditya Acharya、Auguste H. Gezalyan 和 David M. Mount；根据知识共享署名 4.0 国际许可协议授权 莱布尼茨国际信息学会议论文集 Schloss Dagstuhl – Leibniz信息学中心，德国达格施图尔，Dagstuhl出版社

在贝尔特拉米–克莱因模型下，双曲度量是希尔伯特度量的一个特例，此时定义凸体为欧几里得球。

只要所关注的定义域是凸体，且距离对接近边界的程度较为敏感，Hilbert 几何和芬克几何就适用。例如，考虑定义域大小为 m 的离散概率分布。这一空间中的每个点都是一个向量（p_1，……，p_m），其中 $0 \leq p_i \leq 1$ 且 $\sum_{i=1}^{m} p_i = 1$。这些约束意味着这些点位于一个凸定义域内，具体而言，即位于（$m-1$）维概率单纯形中。在这一情境下，欧几里得距离并不十分适用，因为与远离 0 和 1 这两个极值的概率变化相比，欧几里得距离无法敏感地反映非常接近关键值 0 和 1 的概率变化。相比之下，当点接近边界时，希尔伯特距离会非常敏感。Nielsen 和 Sun 提供了实证证据，表明在概率单纯形中，希尔伯特度量与其他广泛使用的距离函数相比，在聚类方面具有竞争力 [24]。Hilbert 度量及相关度量的其他应用还包括通过双曲几何分析网络 [20], 、在深度学习中分析正定矩阵 [14, 21], ，以及基于格的密码系统 [2]。

1.1 相关工作

在欧几里得度量下，通常将支持向量机建模为凸优化问题来求解 [4]。 [7] 的作者是最早研究双曲几何中支持向量机的学者之一。他们证明，在实践中，通过遵循数据的内在几何结构，他们的方法在多个基准测试中的表现显著优于欧几里得分类器。然而，与欧几里得支持向量机的凸优化形式不同，他们的方法需要求解非凸优化问题。因此，若不进行谨慎的初始化，该方法既无法保证得到全局最优解，也无法保证最坏情况下的多项式运行时间。

最近， [1] 的作者提出了一种算法，用于求解与本文相同的多面体支持向量机问题。他们证明，支持向量机问题可以建模为 LP 型问题。该算法的运行时间与点的数量呈线性关系，但会随着定义该度量的多胞体的总复杂度呈多项式增长。遗憾的是，该多胞体的总复杂度通常会随空间维数呈指数增长。因此，他们的方法仍不适用于高维场景，而机器学习中的大多数应用都处于高维空间。相比之下，本文算法的运行时间关于维数呈多项式增长。

1.2 我们的贡献

本文从 Hilbert 几何的另一种视角出发，为希尔伯特支持向量机问题开发了一种高效算法。我们方法的主要特点是，其运行时间仅随空间的维数呈多项式增长。此外，我们的算法可归约为线性规划，而目前已有高效且实用的软件解决方案。由于我们依赖线性规划的数值解，因此需要假设表示输入所需的比特数是有界的，并且将输出计算到用户指定的近似误差范围内。下面的定理将形式化地表述我们的主要结果。

▶**定理 1.** 考虑 $\mathbb{R}^d$ 中由 m 个超平面定义的多胞体 Ω，以及包含在 Ω 内部的两个线性可分点集 P^+ 和 P^-，其总大小为 n。进一步假设，所有定义坐标和系数均可表示为 B 位有理数。那么，存在一种运行时间为 $O(\textit{poly}\,(n, m, d, B, \log(1/\varepsilon))$ 的算法，可以计算出一个分离超平面，其在由Ω 定义的希尔伯特度量下的间隔与最优值相差不超过 ε 的加性误差。

我们方法的显著特点是对空间维数呈多项式依赖，这也是我们所设想的应用中的一个关键假设。我们还将结果推广到密切相关的芬克几何（定义见第 2.1 节）。

除了线性可分的情形，我们还为支持向量机问题提供了一种软间隔 形式，该形式适用于两个点集不可线性分的情况（第 4.1 节）。最后，我们讨论了希尔伯特度量嵌入赋范空间时的基于最近邻的分类器（第 4.2 节）。请注意，与欧几里得情形中平分线是超平面不同，希尔伯特平分线并非超平面。因此，这个问题无法归约为支持向量机问题。

2 数学预备知识

本节将给出一些定义和事实，供全文使用。我们将使用不同的字体，具体而言是 p 和 $\mathbf{p}$，以区分点 $p \in \mathbb{R}^d$ 及其对应的 d 维列向量 $\mathbf{p}$。我们用一对 $(\mathbf{w}\ c) \in \mathbb{R}^d \times \mathbb{R}$ 表示 $(d-1)$ 维平坦集（即超平面） L 在 $\mathbb{R}^d$ 中的形式，该超平面由函数 $f_L(x) = \mathbf{w}^\top \mathbf{x} + c$表示。超平面将空间划分为三个集合 L^-、 L、 L^+，其中点 $\mathbf{x} \in \mathbb{R}^d$ 所属的集合取决于 $f_L(x)$ 分别为 < 0、$= 0$ 还是 > 0。

给定一组 m 超平面，$\mathcal{L} = \{L_i : (\mathbf{w}_i, c_i)\}_{i=1}^m$ 令 $\Omega = \Omega(\mathcal{L})$ 表示由正半空间 L_i^+ 界定的空间的闭包，即 $\Omega = \mathrm{cl}\big(\bigcap_{i=1}^m L_i^+\big)$。令 $\partial\Omega$ 表示 Ω 的边界。全文中，我们假设 Ω 有界且具有满维数。给定一个非负整数 m，令 $[m]$ 表示指标集$\{1, \ldots, m\}$。

给定两个点 $p, q \in \mathrm{int}(\Omega)$，令 pq 表示经过这两点的 Ω 的弦，即 Ω 与经过 p 和 q的直线的交集。令 $\overleftrightarrow{pq}$ 表示经过它们的直线，令 $\overrightarrow{pq}$ 表示从 p 出发并经过 q的射线。

给定 $p, q \in \mathbb{R}^d$，令 $d(p, q)$ 表示它们之间的欧几里得距离。给定一个超平面 L，令$d(p, L)$ 表示从 p 到 L的最小欧几里得（垂直）距离。

2.1 希尔伯特几何

本节介绍芬克距离和希尔伯特距离。给定一个凸体 Ω 以及其中任意两个点 $p, q \in \mathrm{int}(\Omega)$，如果 $p = q$，则所有距离都定义为零。否则，令p'和 q'表示弦 pq的端点，其顺序为 $\langle p', p, q, q' \rangle$（见图1）。芬克距离和希尔伯特距离定义如下:

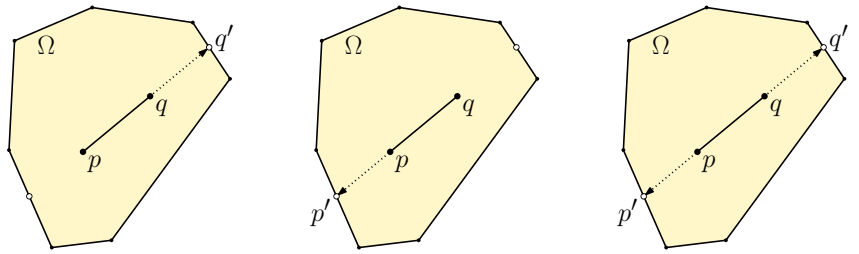
$$\text{Funk:} \quad d_\Omega^F(p,q) := \ln \frac{\|p-q'\|}{\|q-q'\|} \tag{1}$$

$$\text{Reverse Funk:} \quad d_\Omega^{rF}(p,q) := d_\Omega^F(q,p) = \ln \frac{\|q-p'\|}{\|p-p'\|} \tag{2}$$

$$\text{Hilbert:} \quad d_\Omega^H(p,q) := \frac{d_\Omega^F(p,q) + d_\Omega^{rF}(p,q)}{2} = \frac{1}{2} \ln \frac{\|q-p'\|}{\|p-p'\|} \frac{\|p-q'\|}{\|q-q'\|}. \tag{3}$$

直观地说， $d_\Omega^F(p, q)$ 衡量 q 相对于 q 之外的边界距离 p 有多远，而 $d_\Omega^{rF}(p, q)$ 则恰好相反。希尔伯特距离 $d_\Omega^H(p, q)$ 通过取二者的算术平均使其对称化。

这些距离函数均为非负，并满足三角不等式 [26]。希尔伯特距离在 $\mathrm{int}(\Omega)$ 上定义了一种度量。芬克度量和反向芬克度量是不对称的，因此在 $\mathrm{int}(\Omega)$ 上定义了弱度量。注意，当任一点趋近 Ω 的边界时，希尔伯特距离趋近于 ∞。希尔伯特距离在可逆射影变换下保持不变。



(a) (b) (c)

图1 Ω 中 p 与 q 之间的 (a) 正向芬克度量、(b) 反向芬克度量以及 (c) 希尔伯特距离。

2.2 比尔霍夫公式与度量球

比尔霍夫 [5]为这些度量提供了一种等价的替代定义，这种定义更适合我们的方法。

▶**命题2** (比尔霍夫公式)。 给定一组 m 超平面 $L = \{L_i : (\mathbf{w}_i, c_i)\}_{i=1}^m$，那么对于任意一对点 $p, q \in \text{int}(\Omega(L))$:

$$
\begin{aligned}
d_\Omega^F(p,q) &= \max_{i\in[m]} \log \frac{d(p,L_i)}{d(q,L_i)} = \max_{i\in[m]} \log \frac{\mathbf{w}_i^\top \mathbf{p} + c_i}{\mathbf{w}_i^\top \mathbf{q} + c_i} \\
d_\Omega^{rF}(p,q) &= \max_{i\in[m]} \log \frac{d(q,L_i)}{d(p,L_i)} = \max_{i\in[m]} \log \frac{\mathbf{w}_i^\top \mathbf{q} + c_i}{\mathbf{w}_i^\top \mathbf{p} + c_i} \\
d_\Omega^H(p,q) &= \frac{1}{2} \max_{j,k\in[m]} \log \frac{d(p,L_j)}{d(q,L_j)} \frac{d(q,L_k)}{d(p,L_k)} = \frac{1}{2} \max_{j,k\in[m]} \log \left(\frac{\mathbf{w}_j^\top \mathbf{p} + c_j}{\mathbf{w}_j^\top \mathbf{q} + c_j} \frac{\mathbf{w}_k^\top \mathbf{q} + c_k}{\mathbf{w}_k^\top \mathbf{p} + c_k} \right).
\end{aligned}
$$

这是由 Ω 为凸集这一事实推得的；注意，如果 $\overrightarrow{pq}$ 与 $\partial\Omega$ 在 q'处相交，且该点位于支撑超平面 L_k上，那么根据相似三角形

$$
\frac{d(p,q')}{d(q,q')} = \frac{d(p,L_i s)}{d(q,L_i)}.
$$

直观而言，这提供了一种定义度量的方法，无须显式计算直线 $\overleftrightarrow{pq}$ 与 $\partial\Omega$ 的交点。该命题的一个推论是，利用非线性映射函数 [22]，度量可以化简为 $\mathbb{R}^t$ 上某个有限维数 t空间中的上确界范数。另一个有用的推论是，它可用于在高维空间中构造度量球。

令 d_Ω^* 表示上述任一度量，给定点 $p \in \text{int}(\Omega)$ 和标量 $r \geq 0$，相应的度量球按标准方式定义。

$B_\Omega^*(p, r) := \{q \in \Omega : d_\Omega^*(p, q) \leq r\}$。

众所周知，关于点 p 的Funk度量是将Ω以 p 为中心按一个适当比例缩放得到的，该比例取决于半径。这可以直接由命题2推导得出：

$$\begin{aligned} B_\Omega^F(p,r) &= \{q \in \Omega : d_\Omega^F(p,q) \le r\} \\ &= \left\{q \in \Omega : \max_{i\in[m]} \log \frac{\mathbf{w}_i^\top \mathbf{p} + c_i}{\mathbf{w}_i^\top \mathbf{q} + c_i} \le r\right\} \\ &= \left\{q \in \Omega : \log \frac{\mathbf{w}_i^\top \mathbf{p} + c_i}{\mathbf{w}_i^\top \mathbf{q} + c_i} \le r,\ \forall i \in [m]\right\} \\ &= \{q \in \Omega : 0 \le \mathbf{w}_i^\top \mathbf{q} + c_i - e^{-r}(\mathbf{w}_i^\top \mathbf{p} + c_i),\ \forall i \in [m]\} \\ &= \bigcap_{i\in[m]} \{q \in \Omega : 0 \le \mathbf{w}_i^\top \mathbf{q} + c_i - e^{-r}(\mathbf{w}_i^\top \mathbf{p} + c_i)\}. \end{aligned} \tag{4}$$

固定 p 和 r，令 $c_i^* := c_i - e^{-r}(\mathbf{w}_i^\top \mathbf{p} + c_i)$，并令 S_i^F 为由 $(\mathbf{w}_i, c_i^*)$ 表示的超平面（见图 2(a)）。注意，S_i^F 与 L_i 平行。我们有如下结论：

▶ **引理3** (芬克球)。给定一组 m 超平面 L，如命题2所述，芬克球 $B_{\Omega(\mathcal{L})}^F(p,r)$ 是一个多胞体，其边界由 m 超平面 S_i^F: $(\mathbf{w}_i, c_i^*)$ 界定，其中 $i \in [m]$。

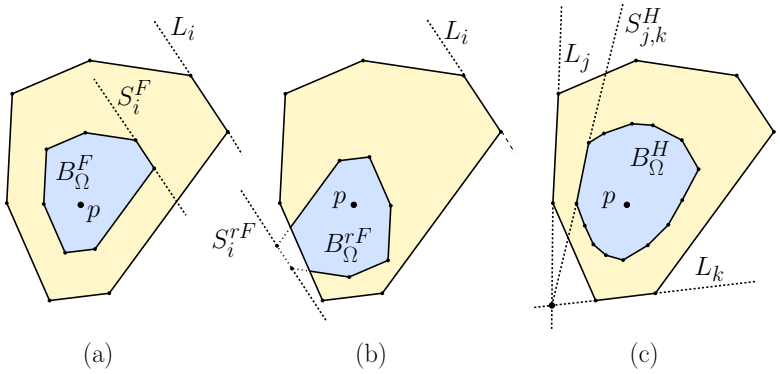


图2 定义球及其支撑超平面：(a) 芬克，(b) 反向芬克，(c) 希尔伯特。

反向芬克球也可以证明有类似的结果（见图 2(b)）。

接下来，我们使用命题2证明希尔伯特球是凸多胞体，但其边界面的数量增加为原来的平方。按照式（4）中的相同推理，我们有

$$\begin{aligned} B_\Omega^H(p,r) &= \{q \in \Omega : d_\Omega^H(p,q) \le r\} \\ &= \left\{q \in \Omega : \frac{1}{2} \max_{j,k\in[m]} \log \left(\frac{\mathbf{w}_j^\top \mathbf{p} + c_j}{\mathbf{w}_j^\top \mathbf{q} + c_j} \frac{\mathbf{w}_k^\top \mathbf{q} + c_k}{\mathbf{w}_k^\top \mathbf{p} + c_k}\right) \le r\right\} \\ &= \bigcap_{j,k\in[m]} \left\{q \in \Omega : \left(\mathbf{w}_j^\top \mathbf{p} + c_j\right)\left(\mathbf{w}_k^\top \mathbf{q} + c_k\right) \le e^{2r}\left(\mathbf{w}_k^\top \mathbf{p} + c_k\right)\left(\mathbf{w}_j^\top \mathbf{q} + c_j\right)\right\}. \end{aligned}$$

固定 p 和 r，令 $\alpha_{j,k} := e^{2r}\left(\mathbf{w}_k^\top \mathbf{p} + c_k\right)$ 和 $\alpha'_{j,k} := \left(\mathbf{w}_j^\top \mathbf{p} + c_j\right)$。我们有

$$B_\Omega^H(p,r) = \bigcap_{j,k\in[m]} \left\{q \in \Omega : 0 \le \left(\alpha_{j,k}\mathbf{w}_j - \alpha'_{j,k}\mathbf{w}_k\right)^\top \mathbf{q} + \left(\alpha_{j,k}c_j - \alpha'_{j,k}c_k\right)\right\}. \tag{5}$$

令 $\mathbf{w}^*_{j,k} = (\alpha_{j,k}\mathbf{w}_j - \alpha'_{j,k}\mathbf{w}_k)$，以及 $c^*_{j,k} = (\alpha_{j,k}c_j - \alpha'_{j,k}c_k)$。令 $S^H_{j,k}$ 为由 $(\mathbf{w}^*_{j,k}, c^*_{j,k})$ 表示的超平面。注意，$S^H_{j,k}$ 是 L_j和 L_k的线性组合，因此 $S^H_{j,k}$ 经过 L_j 和 L_k 的交点（见图 2(c)）。综上，我们有如下结论：

▶**引理 4** (希尔伯特球)。给定一组如命题2中所述的 m 个超平面 L，希尔伯特球 $B^H_{\Omega(\mathcal{L})}(p, r)$ 是一个多胞体，其边界由至多 $m(m-1)$ 个超平面构成：$S^H_{j,k}$，其中 $j, k \in [m]$。

3 希尔伯特支持向量机

本节介绍我们在希尔伯特几何中解决支持向量机问题的方法。回顾一下，我们的目标是计算一个最大间隔超平面（相对于希尔伯特距离），将两个点集 P^+和 P^-分开，这两个点集的总大小为 n。

给定指标集 $I = \{1, \ldots\ldots, n\}$，以及将 I划分为 I^+和 I^-的一个分割，即 $I^+ \cup I^- = I$， $I^+ \cap I^- = \emptyset$，并给定相应的 d维点集： $P^{\pm} = \{p_i : i \in I\}$， $P^+ = \{p_i : i \in I^+\}$， $P^- = \{p_i : i \in I^-\}$。在本节中，我们始终假设 P^+和 P^- 线性可分，也就是说，存在一个超平面 K 使得 $P^+ \subset K^+$，且$P^- \subset K^-$。（在第 4.1 节中，我们将讨论一般情形。）

定义点 $p \in \Omega$ 到超平面 L 的 Hilbert 距离为 $d^H_\Omega(L, p) = \min_{x \in L} d^H_\Omega(x, p)$。我们将 P^+ 和 P^- 相对于分离超平面 L 的间隔 定义为 $\min_{p \in P^{\pm}} d^H_\Omega(L, p)$，即任意点到 L 的最小距离。求任意分离超平面的最大间隔，可以形式化地表示为

$$\max\{\gamma : \exists K \text{ such that } P^+ \subset K^+, P^- \subset K^-, \text{ and } d^H_\Omega(K, p) \geq \gamma, \forall p \in P^{\pm}\}. \tag{6}$$

我们的算法将得到最优解的绝对近似。具体而言，给定一个近似参数 $\varepsilon > 0$，我们的算法将生成一个分离超平面，其间隔至少为 $\gamma - \varepsilon$，其中 γ 是最优间隔。

假设 K 由数对 $(\mathbf{w}\ c)$ 表示，则这可以等价地表述为以下优化问题：

$$\begin{aligned} \text{maximize:} \quad & \gamma \\ \text{subject to:} \quad & \mathbf{w}^\top \mathbf{p} + c > 0, \quad \forall p \in P^+ \\ & \mathbf{w}^\top \mathbf{p} + c < 0, \quad \forall p \in P^- \\ & d^H_\Omega(K, p) \geq \gamma, \quad \forall p \in P^{\pm}, \quad K : (\mathbf{w}, c). \end{aligned} \tag{7}$$

如何求解 Opt (7) 并非一目了然。一个难点在于，如何为 $d^H_\Omega(K, p)$（点到超平面的希尔伯特距离）找到合理的表示方法。早先一项关于多胞体希尔伯特几何中支持向量机的研究 [1], 解决了这个问题，但其中给出的表示规模取决于希尔伯特球顶点的数量，而顶点数量会随维数呈指数增长。其他在双曲空间中构造支持向量机问题的方法通常仿照欧几里得情形，将其表述为非凸优化问题，且没有关于运行时间的理论保证 [7]。然而，利用多胞体希尔伯特度量的几何性质，并采用命题2中介绍的度量的其他视角，我们能够显著改进以往的尝试。

3.1 总体算法

我们假设每个输入参数都表示为有理数，即两个非零 B 位整数之比。我们将这种表示称为 B 位有理数。每个 n 点 $p \in P^{\pm}$ 均表示为一个 d 维向量 $\mathbf{p}$，其坐标为 B 位有理数。定义域 Ω 由一组 $\mathcal{L}$ 个 m 超平面定义，每个超平面的 $(\mathbf{w}\ c)$ 表示均由一串 $d+1$ B 位有理数组成。我们称由此得到的支持向量机问题实例具有 规模参数 $(d,\ n,\ m,\ B)$。此类实例的总位复杂度 为 $O(d(n+m)B)$。

在描述总体算法之前，我们首先根据输入的位数大小，建立关于 γ 最大间隔的一个上界。令 K 为相应的分离超平面，使得 γ 达到最大值。注意，对于任意 $p_+ \in P^+$ 和 $p_- \in P^-$，都有 $d_\Omega^H(p_-,\ p_+) \geq d_\Omega^H(p_-,\ K) + d_\Omega^H(p_+,\ K) \geq 2\gamma$。现在，令连接 p_- 和 p_+的直线与 $\partial\Omega$ 相交于支撑超平面 L_-：$(\mathbf{w}_-,\ c_-)$，以及 L_+：$(\mathbf{w}_+,\ c_+)$。因此，

$$d_\Omega^H(p_-, p_+) \ = \ \frac{1}{2}\log\frac{d(p_+, L_-)\, d(p_-, L_+)}{d(p_+, L_+)\, d(p_-, L_-)} \ = \ \frac{1}{2}\log\frac{\left(\mathbf{w}_-^\top \mathbf{p}_+ + c_-\right)\left(\mathbf{w}_+^\top \mathbf{p}_- + c_+\right)}{\left(\mathbf{w}_-^\top \mathbf{p}_- + c_-\right)\left(\mathbf{w}_+^\top \mathbf{p}_+ + c_+\right)}.$$

考虑分子中的任意一项，例如 $(\mathbf{w}_-^\top \mathbf{p}_+ + c_-)$。现在，$\mathbf{w}_-$ 和 p_+均由 d B位有理数给出，因此它们的内积的绝对值至多为 $d \cdot 2^{2B}$。因此，整个项 $\mathbf{w}_-^\top \mathbf{p}_+ + c_-$ 的绝对值至多为 $d \cdot 2^{2B} + 2^B$。回忆一下，希尔伯特距离是比值的对数（相差常数因子），$d_\Omega^H(p_-,\ p_+)$ 至多为 $(8B + 4\log\ d)\log\ 2 = O(B + \log\ d)$，这意味着 $\gamma \leq M \in O(B + \log d)$。因此，我们有：

▶ **引理 5**。 给定一个规模参数为(d, n, m, B)的支持向量机实例，任意超平面的分离间隔γ至多为 $M(d, n, m, B) = O(B + \text{对数}\ d)$。

总体算法如下。由于我们关注的是精度达到 ε 的最优间隔，因此我们检查 P^+ 和 P^- 是否能以至少 r 的间隔分离，其中 r 是取自以下范围的一个元素： $\Gamma = [\varepsilon, 2\varepsilon, \ldots, M]$，其中 $M = M(d, n, m, B)$。要检查特定值 r 下的分离情况，我们将进行一次（近似的）基于线性规划的可行性测试。我们将此测试用于范围 Γ 上的参数化二分搜索。这意味着我们只需进行 $O(\log B + \log\log\ d + \log(1/\varepsilon))$ 次可行性测试。接下来我们介绍这个基于线性规划的子程序。

3.2 线性规划可行性

我们考虑这样一个问题：判断是否能够达到给定的间隔γ，误差不超过 ε 的加性误差。对于特定的间隔 r，回忆一下，$B_\Omega^H(p,\ r)$ 是半径为 r、以 p 为中心的 Hilbert 球。根据引理 4，$B_\Omega^H(p,\ r)$ 是一个由$m(m-1)$ 个超平面界定的多胞体。（在第 3.3 节中，我们将说明如何高效地构造它们。）令$\mathcal{B}(r)^+ = \{B_\Omega^H(p^+,\ r) \mid p^+ \in P^+\}$表示以正标记点为中心、半径为 r 的 Hilbert 球的集合，并类似地定义 P^- 对应的$\mathcal{B}(r)^-$。

γ 的可行性可归结为判断球的集合 $\mathcal{B}(r)^+$ 和 $\mathcal{B}(r)^-$ 是否线性可分。对于 $i \in [n]$，令 $\mathbf{A}(r,\ i)$ 为一个 $m(m-1)$ 矩阵，并令 $\mathbf{b}(r,\ i)$ 为一个 $m(m-1) \times 1$ 向量，使得

$$\mathbf{x} \in B_\Omega^H(p_i, r) \iff \mathbf{A}(r,i)\mathbf{x} + \mathbf{b}(r,i) \geq 0.$$

我们可以根据引理 4 中描述的超平面 $S_{j,k}^H$：$(\mathbf{w}_{j,k}^*, c_{j,k}^*)$ 构造 $\mathbf{A}(r,i)$ 和 $\mathbf{b}(r,i)$。特别地，对于 $t \in [m(m-1)]$,，如果 $(\mathbf{w}_t^*, c_t^*)$ 表示 $B_\Omega^H(p_i, r)$ 的第 t 个支撑超平面，那么 $\mathbf{w}_t^*$就是 $\mathbf{A}(r,i)$ 的第 t 行，而 c_t^* 是 $\mathbf{b}(r,i)$ 的第 t 个元素。

因此，判断$\mathcal{B}(r)^+$ 和 $\mathcal{B}(r)^-$ 是否线性可分，等价于以下判定问题：

$$
\begin{aligned}
&\exists K : (\mathbf{w}, c), \quad \text{such that} \\
&\quad \forall p_i \in P^+, \\
&\qquad \forall \mathbf{x} : \mathbf{A}(r,i)\mathbf{x} + \mathbf{b}(r,i) \geq 0, \quad \mathbf{w}^\top \mathbf{x} + c \geq 0, \qquad (8) \\
&\quad \forall p_i \in P^-, \\
&\qquad \forall \mathbf{x} : \mathbf{A}(r,i)\mathbf{x} + \mathbf{b}(r,i) \geq 0, \quad \mathbf{w}^\top \mathbf{x} + c \leq 0. \qquad (9)
\end{aligned}
$$

固定 r 和 i，令 $\mathbf{A} = \mathbf{A}(r, i)$ 和 $\mathbf{b} = \mathbf{b}(r, i)$。式（8）可改写为：

$$\min_{\mathbf{x}:\, \mathbf{A}\mathbf{x}+\mathbf{b}\geq 0} \mathbf{w}^\top \mathbf{x} + c \;\geq\; 0.$$

根据标准的线性规划对偶性 [23]，我们有

$$\min_{\mathbf{x}:\, \mathbf{A}\mathbf{x}+\mathbf{b}\geq 0} \mathbf{w}^\top \mathbf{x} \quad = \quad \max_{\mathbf{y}:\, \mathbf{A}^\top \mathbf{y}=\mathbf{w},\, \mathbf{y}\geq \mathbf{0}} -\mathbf{b}^\top \mathbf{y}.$$

因此，我们可以将式（8）简化如下：

$$
\begin{aligned}
&\forall \mathbf{x} : \mathbf{A}\mathbf{x} + \mathbf{b} \geq 0, \quad \mathbf{w}^\top \mathbf{x} + c \;\geq\; 0 \\
&\quad \Longleftrightarrow \min_{\mathbf{x}:\, \mathbf{A}\mathbf{x}+\mathbf{b}\geq 0} \mathbf{w}^\top \mathbf{x} + c \;\geq\; 0 \\
&\quad \Longleftrightarrow \max_{\mathbf{y}:\, \mathbf{A}^\top \mathbf{y}=\mathbf{w},\, \mathbf{y}\geq \mathbf{0}} -\mathbf{b}^\top \mathbf{y} + c \;\geq\; 0 \\
&\quad \Longleftrightarrow \exists \mathbf{y}(i) : \; \mathbf{A}(r,i)^\top \mathbf{y}(i) \;=\; \mathbf{w}, \;\; \mathbf{y}(i) \geq \mathbf{0}, \;\; \mathbf{b}(r,i)^\top \mathbf{y}(i) \leq c.
\end{aligned}
$$

类似地，式（9）可简化为

$$\exists \mathbf{y}(i) : \; \mathbf{A}(r,i)^\top \mathbf{y}(i) \;=\; \mathbf{w}, \;\; \mathbf{y}(i) \geq \mathbf{0}, \;\; \mathbf{b}(r,i)^\top \mathbf{y}(i) \geq c.$$

因此，要确定集合 $\mathcal{B}(r)^+$ 和 $\mathcal{B}(r)^-$ 是否线性可分，我们求解关于参数$\mathbf{w}_{d\times 1}$、 c 和$\mathbf{y}(i)$ $_{m(m-1)\times 1}$（其中$i \in [n]$）的以下线性规划可行性问题。

$$
\begin{aligned}
\text{最小化:} \quad & 0 \\
\text{约束条件:} \quad & \mathbf{b}(r,i)^\top \mathbf{y}(i) \;\leq\; c, \quad \forall i \in I^+ \\
& \mathbf{b}(r,i)^\top \mathbf{y}(i) \;\geq\; c, \quad \forall i \in I^- \\
& \mathbf{A}(r,i)^\top \mathbf{y}(i) \;=\; \mathbf{w}, \quad \forall i \in I \\
& \mathbf{y}(i) \;\geq\; \mathbf{0}, \quad \forall i \in I \\
& \mathbf{1}^\top \mathbf{w} \;=\; 1. \qquad (10)
\end{aligned}
$$

We 添加最后一个约束条件，以避免所有参数均为零这一平凡解 .

▶ 注 6（直观描述）。通过观察以下内容，可以直观地理解 Opt. (10)。$\mathcal{B}(r)^+$ 和$\mathcal{B}(r)^-$ 线性可分，意味着存在一个超平面K：（$\mathbf{w}$ c），使得对于每个 $B^+ \in \mathcal{B}(r)^+$，都存在一个与 K 平行且位于其上方的超平面 K_{B^+}，并且该超平面与 B^+中的某个球相切。类似地，对于每个 $B^- \in \mathcal{B}(r)^-$，都存在一个与 K平行且位于其下方的超平面 K_{B^-}，并且该超平面与 B^-中的某个球相切（见图3）。

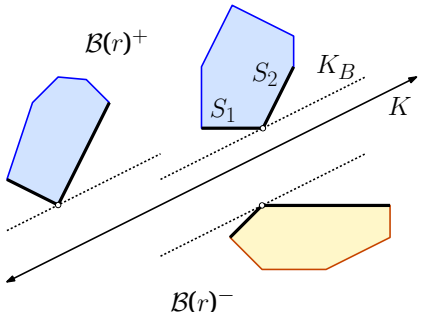


图3 对偶的直观描述：K 将$\mathcal{B}(r)^+$ 与$\mathcal{B}(r)^-$分隔开。 K_B 是 S_1 和 S_2 的锥组合，同时严格位于 K 上方。

与多胞体 Ω 相切的超平面 L，其支撑超平面为 $L_1, \ldots, L_m$，可以表示为这些支撑超平面的正线性组合（通常称为锥组合）。

优化问题 (10) 中的向量 $\mathbf{y}(i)$ 对应于该锥组合的系数。优化问题 (10) 中的第三个约束保证存在一个与 K 平行的超平面 K_B，且该超平面是每个 $B \in \mathcal{B}(r)^+ \cup \mathcal{B}(r)^-$ 的支撑超平面的锥组合。前两个约束确保 K_B 位于 K 上方或下方，具体取决于 $B \in \mathcal{B}(r)^+$ 还是 $\mathcal{B}(r)^-$。

▶ 备注 7（Funk度量的线性规划）。Funk度量的线性规划可行性表述与优化问题 (10) 类似。下面简要说明所需的修改。根据引理3中的构造，对于 $i \in [n]$，令 $\mathbf{A}(r,i)$ 为一个 $m \times d$ 矩阵，并令 $\mathbf{b}(r,i)$ 为一个 $m \times 1$ 向量，使得

$$\mathbf{x} \in B_\Omega^F(p_i, r) \iff \mathbf{A}(r,i)\mathbf{x} + \mathbf{b}(r,i) \geq 0.$$

对于这种情况，$\mathbf{y}(i)$ 将是一个 $m \times 1$ 向量。其余的分析和表述与希尔伯特度量的情况完全相同。

3.3 构造Hilbert球

本节介绍一种构造希尔伯特球 $B_\Omega^H(p_i, r)$ 的支撑超平面的方法。回顾一下，$\mathbf{L} = \{L_i\}_{i=1}^m$ 表示界定 Ω 的超平面集合。首先，取 Ω 的两个支撑超平面：$L_j : (\mathbf{w}_j, c_j)$ 和 $L_k : (\mathbf{w}_k, c_k)$，$j \neq k$, $j, k \in [m]$。对于 $p_i \in P^{\pm}$，我们找到 $S(i,r)_{j,k} : (\mathbf{u}(i,r)^{(j,k)}, t(i,r)_{j,k})$，它是相对于由 L_j 和 L_k 界定的边界，与 p_i 的 Hilbert 距离为 r 的超平面；使得存在一条直线 ℓ，按此顺序与 L_j, $S(i,r)_{j,k}$, p_i 和 L_k 相交（见图4）。

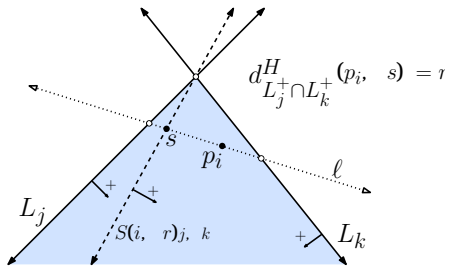


图4 构造 $B_\Omega^H(p_i, r)$。

显然，$\exists\,\alpha：0\le\alpha\le 1$，使得

$$S(i,r)_{j,k} = \alpha L_j - (1-\alpha)L_k. \tag{11}$$

对于 $S(i,\ r)_{j,\ k}, s\in L_j^+\cap L_k^+$ 上的任意点 s，不难看出

$$\frac{d(s,L_k)}{d(s,L_j)} = \frac{\alpha}{1-\alpha}.$$

使用以下内容求解 α：

$$\exp(2\cdot d_\Omega^H(p_i,S(i,r)_{j,k})) = \frac{\mathbf{w}_j^\top\mathbf{p}_i+c_j}{\mathbf{w}_k^\top\mathbf{p}_i+c_k}\cdot\frac{\alpha}{1-\alpha} = \exp(2r). \tag{12}$$

对于 $p_i\in$int(Ω)，$0<r<\infty$。因此，公式（12）可以等价地写成关于 α 的线性方程。由于对于 $j,\ k\in[m], j\neq k, i\in[n], S(i,\ r)_{j,\ k}$ 是 $B_\Omega^H(p_i,\ r)$ 的支撑超平面，我们可以利用公式（11）和公式（12）构造 $B_\Omega^H(p_i,\ r)$，对 $p_i\in P^\pm$ in $O(ndm^2)$ 中的所有时刻均如此。

▶注 8（Funk 球）。我们可以采用类似的方法构造 $B_\Omega^F(p_i,\ r)$，即以 p_i 为中心、半径为 r 的 Funk 球。回顾一下，以 p_i 为中心的 Funk 球是 Ω 以 p_i 为中心的缩放副本。因此，对于 Ω 的支撑超平面 L_j：$(\mathbf{w}_j,\ c_j)\ j\in[m]$，我们有 $S(i,\ r)_j=(\mathbf{w}_j,\ c_j^*(i,\ r))$，它是某个常数 $c_j^*(i,\ r)$ 下 $B_\Omega^F(p_i,\ r)$ 的支撑超平面。我们使用以下方程求解 $c_j^*(i,\ r)$：

$$\exp(d_\Omega^F(p_i,S(i,r)_j)) = \frac{\mathbf{w}_j^\top\mathbf{p}_i+c_j}{\mathbf{w}_j^\top\mathbf{p}_i+c_j^*(i,r)} = \exp(r). \tag{13}$$

与之前一样，对于任意 $r\ge 0$，它都可以表示为关于 $c^*(i,\ r)$ 的线性函数。因此，我们可以在 O 时间内为所有 $p_i\in P^\pm$ 构造 $B_\Omega^F(p_i,\ r)$。

3.4 复杂度分析

回顾一下，为了在误差不超过 ε 的情况下获得 γ，我们需要求解 $O(\log B+\log\log\ d+\log(1/\varepsilon))$ 个形式如 Opt. (10) 的线性规划。我们可以使用任意标准内点法，例如 Khachiyan 椭球法 [18] 或 Karmarkar 方法 [16]，在多项式时间内求解 Opt. (10)；该时间关于变量数、不等式数以及每个变量的位长呈多项式关系。我们首先确定在我们的线性规划可行性问题（Opt. (10)）中编码一个参数所需的位长。

回顾一下，我们在范围 $\Gamma=[\varepsilon,2\varepsilon,\ldots,M]$ 内搜索 r，其中 $M=M(d,\ n,\ m,\ B)=O(B+\log\ d)$。因此，量 e^{2r} 的取值范围为 $\Gamma_e=[e^{2\varepsilon},e^{4\varepsilon},\ldots,e^{2M}]$。我们无需在 Γ 上精确确定 r 的值；就我们的目的而言，只需在 $[r_1,r_2,\ldots,r_{M/\varepsilon}]$ 上搜索 r 即可，其中 $(i-1)\varepsilon\le r_i<i\varepsilon$。

Γ_e 中两个元素之间的最小差值至少为 $e^{4\varepsilon}-e^{2\varepsilon}>\varepsilon$。（可利用 $x=0$ 处 e^x 的泰勒级数验证这一点。）因此，为了表示 Γ_e，我们只需找到一种位表示，使最大值e^{2M} 能够达到 ε 的分辨率。具体而言，可使用 $O(M)$ 位表示整数部分，并使用$O(\log(1/\varepsilon))$ 位表示小数部分。因此，只需将 Γ_e中的元素表示为 $O(M+\log(1/\varepsilon))=O(B+\log d+\log(1/\varepsilon))$ 位有理数。

利用公式（12），我们求解并表示 α，使用 $O(B+\log d+\log(1/\varepsilon))$ 位有理数。因此，利用公式（11），$S(i,\ r)_{j,\ k}$ 的每个参数——其充当界定

超平面，用于 $B_\Omega^H(p_i,\ r)$——也可以使用 $O(B+\log\ d+\log(1/\varepsilon))$ 位有理数表示。

现在，Opt. (10) 中的变量数为 $O((n+d)m^2)$，不等式数为 $O(n(m^2+d))$。因此，对于固定的 r，我们可以在 $O(\text{poly}(n,m,d)(B+\log d+\log(1/\varepsilon))$ 时间内确定 Opt. (10) 是否可行。由于我们要解 $O(\log\ B+\log\log\ d+\log(1/\varepsilon))$ 个这样的线性规划，总运行时间为 $O(\text{poly}(n,m,d,B,\log(1/\varepsilon))$。这证明了定理 1 中的主要结果。回想一下，输入位长度为 $d(n+m)B$，因此总运行时间也关于输入位长度和所需精度的对数呈多项式。

4 扩展

本节将探讨希尔伯特几何中二元分类和支持向量机问题的两种扩展。第一种是软间隔分类器，适用于点集线性不可分的情形；第二种是基于最近邻的二元分类器。

4.1 软间隔分类器

到目前为止，我们假设 P^+ 和 P^- 线性可分。如果它们不可分，一种常见做法是考虑一个超平面分类器，并对训练集中点的误分类施加适当的惩罚。本节将提出一种此类“软”分类器的形式化表述。

我们修改线性规划可行性问题 Opt (10)，引入以松弛变量形式表示的惩罚。重申注 6 中的观察，我们注意到，Opt (10) 的前两个约束确保条件成立：对于 $p\in P^+$, $B_\Omega^H(p,\ r)$，其完全位于分隔超平面 K: $(\mathbf{w}\ c)$ 上方。现在，我们只要求对于某个正类 ξ，$B_\Omega^H(p,\ r)$ 位于 K': $(\mathbf{w}\ c+\xi)$ 上方。此处，ξ 作为 $B_\Omega^H(p,\ r)$ 非适当分离的惩罚。我们同样为每个点分别设定惩罚。

接下来，我们希望满足以下理想性质：离 Ω 的边界越近的点，被误分类时所受的惩罚越重。原因在于，对于固定的点 q,$d_\Omega^H(p_i,\ q)$ 与 p_i 到边界的距离成反比。因此，我们将与 p_i 相关联的惩罚乘以 ω_i，其中 ω_i 定义为该点到边界距离的倒数：

$$\omega_i=\frac{1}{\min_{j\in[m]}d_\Omega^H(L_j,p_i)}$$

最后，对于固定的希尔伯特半径 r，我们提出以下关于参数 $\mathbf{w}$、c 和 ξ_i、$\mathbf{y}(i)$（$i\in[n]$）的线性规划，以最小化总加权惩罚：

$$\begin{aligned}
\text{Minimize:}\qquad \Xi_r &= \sum_{i\in[n]}\omega_i\xi_i\\
\text{约束条件:}\qquad \mathbf{b}(r,i)^\top\mathbf{y}(i) &\le c+\xi_i,\quad \forall i\in I^+\\
\mathbf{b}(r,i)^\top\mathbf{y}(i) &\ge c-\xi_i,\quad \forall i\in I^-\\
\mathbf{A}(r,i)^\top\mathbf{y}(i) &= \mathbf{w},\quad \forall i\in I\\
\mathbf{y}(i) &\ge \mathbf{0},\quad \forall i\in I\\
\mathbf{1}^\top\mathbf{w} &= 1\\
\xi_i &\ge 0,\quad \forall i\in I.
\end{aligned}\tag{14}$$

设由最优化问题（14）得到的最优值为 Ξ_r，相应的分离超平面为；$K_r=(\mathbf{w}\ c)$。回想一下，r 在 $\Gamma_r=[r_1,r_2,\ldots,r_{M/\varepsilon}]$ 上取值，其中 r_i 的

是满足（$i-1$）$\varepsilon \le r_i < i\varepsilon$ 的任意值，且 $M = O(B + \log d)$。为找到合适的分类器 K^*，我们取

$$K^* = K_s, \qquad \text{where} \qquad s = \arg\max_{i\in[M/\varepsilon]} \left(r_i - C\,\frac{\Xi_{r_i}}{n} \right).$$

请注意，C 是用户定义的常数，是与非适当分离相关联的权重。该值越大，越倾向于适当分离；该值越小，则越倾向于获得更大的间隔。

这种方法的一个小缺点是，仅在范围 Γ_r内进行二分搜索已不足够。相反，我们需要进行网格搜索，也就是说，最坏情况下要针对 $r \in \Gamma_r$ 的每个取值求解 Opt. (14)。这意味着软间隔分类器的总体运行时间为 $O(\text{poly}(n,\ m,\ d,\ B,\ 1/\varepsilon))$。

4.2 基于最近邻的分类器

本节中，我们将探讨一种在 Hilbert 几何中进行二元分类的替代方法。我们不使用分离超平面，而是选取两个站点 q^+ 和 q^-，并根据点 p 离这两个站点中哪一个更近来对其进行分类。在欧几里得情形下，两点之间的等距线是一个超平面，因此这种方法与支持向量机并无不同。然而，在 Hilbert 几何中，平分线通常不是超平面。（ [12], 中分析了二维情形下 Hilbert 平分线的结构，并指出它们是分段二次曲线。）

和之前一样，我们假设希尔伯特几何是相对于由 m 个支撑超平面$\mathcal{L} = \{L_i\}_{i=1}^m$定义的 $\mathbb{R}^d$ 中的多胞体 Ω 而定义的。Vernicos 扩展了 de la Harpe 的结果，提出了一个优雅的等距映射，将多面体希尔伯特几何（Ω d_Ω^H）映射到有限维赋范向量空间 [32]。具体而言，他给出了一个等距映射f_Ω： (Ω d_Ω^H) $\to (\mathbb{R}^{m-1}, ||\cdot||_{\Sigma_{2m}})$，其中 $||_{\Sigma_{2m}}$上的范数· 由 $\mathbb{R}^{m-1}$中的一个正则多胞体 Σ定义，该正则多胞体由 $2m$ 个面界定。换言之，嵌入空间是 $\mathbb{R}^{m-1}$ ，其单位球为 Σ_{2m} 。

尽管人们可能会想在嵌入空间中求解支持向量机问题，但^结果并不理想。由于 f_Ω 是非线性的，嵌入空间中的线性分类器 K 无法转化为输入空间中的简单分类器。此外，由于 f_Ω 不是^双射，因此难以确定 K 与 Ω 的像 f_Ω(Ω)相交后会有多复杂。

相反，我们提出一种在嵌入空间中简单的基于最近邻的分类器。给定 Ω，令 q_i、Q^+ 和 Q^- 分别表示 p_i、P^+ 和 P^- 在嵌入 f_Ω 下的映像。（对所有 $i \in [n]$，q_i 均可通过求解 m 个线性方程，在 $O(m^2)$ 时间内计算得到。）我们分别为 Q^+ 和 Q^- 找到两个代表中心 c^+ 和 c^-，使得以下表达式最大化：

$$\beta = \max_{c^+,c^-\in\mathbb{R}^{m-1}} \left(\min_{q^+\in Q^+} \frac{\min_{q^-\in Q^-} \|\mathbf{q}^+ - \mathbf{q}^-\|_\Sigma}{\|\mathbf{q}^+ - \mathbf{c}^+\|_\Sigma}, \min_{q^-\in Q^-} \frac{\min_{q^+\in Q^+} \|\mathbf{q}^- - \mathbf{q}^+\|_\Sigma}{\|\mathbf{q}^- - \mathbf{c}^-\|_\Sigma} \right) \tag{15}$$

换句话说，β 是这样一个值：对于任意点 q，它到相反类别中最近邻的距离至少是到其代表点距离的 β 倍。这意味着，使用近似因子至多为 β 的近似最近邻算法即可正确分类这些点。

这可以改写为关于变量 $\mathbf{c}^+$、$\mathbf{c}^-$ 和 $1/\beta$ 的一个优化问题

$$\begin{aligned} &\text{最小化:} && \frac{1}{\beta} \\ &\text{约束条件:} && \|\mathbf{q}^+ - \mathbf{c}^+\|_\Sigma \;\le\; \frac{1}{\beta}\|\mathbf{q}^+ - \mathbf{q}^-\|_\Sigma, \quad \forall q^+ \in Q^+, \forall q^- \in Q^- \\ & && \|\mathbf{q}^- - \mathbf{c}^-\|_\Sigma \;\le\; \frac{1}{\beta}\|\mathbf{q}^- - \mathbf{q}^+\|_\Sigma, \quad \forall q^+ \in Q^+, \forall q^- \in Q^- \end{aligned} \tag{16}$$

我们注意到，上述优化中的 $\|\mathbf{q}^+ - \mathbf{q}^-\|_\Sigma$ 是一个预先确定的常数。若令 $W = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{2m}\}$ 表示确定范数多胞体 Σ 的方向，则 $\|\mathbf{x}\|_\Sigma \le t$ 的条件可以表示为一组线性不等式：$\mathbf{w}_i^\top \mathbf{x} \le t$，其中 $i \in [2m]$。因此，优化问题（16）是一个包含 $O(m)$ 个变量和 $O(nm)$ 个不等式的线性规划。

参考文献

1 A. Acharya, A. H. Gezalyan, J. Vanecek, D. M. Mount, and S. Arya. 希尔伯特几何中的支持向量机。载于第19届国际算法与数据结构研讨会论文集，第3:1–3:20页，2025年。doi:10.4230/LIPIcs.WADS.2025.3。2 S. Arya, G. D. da Fonseca, and D. M. Mount. 经济凸覆盖及其应用。SIAM计算杂志，53(4):1002–1038，2024年。doi:10.1137/23M1568351。3 A. Ben-Hur, C. S. Ong, Sören Sonnenburg, Bernhard Schölkopf, and G. Rätsch. 用于计算生物学的支持向量机与核函数。PLoS计算生物学，4(10):e1000173，2008年。doi:10.1371/journal.pcbi.1000173。4 K. P. Bennett and C. Campbell. 支持向量机：炒作还是值得喝彩？SIGKDD探索通讯，2，2000年。doi:10.1145/380995.380999。5 G. Birkhoff. Jentzsch定理的推广。美国数学会汇刊，85:219–227，1957年。6 C. M. Bishop. 模式识别与机器学习。施普林格，2016年。doi:10.1007/978-0-387-45528-0。7 H. Cho, B. DeMeo, J. Peng, and B. Berger. 双曲空间中的大间隔分类。载于第22届国际人工智能与统计学会议论文集，第1832–1840页。PMLR，2019年。doi:10.48550/arXiv.1806.00437。8 C. Cortes and V. Vapnik. 支持向量网络。机器学习，20:273–297，1995年。doi:10.1007/BF00994018。9 N. Cristianini and J. Shawe-Taylor.《支持向量机及其他基于核的学习方法导论》。剑桥大学出版社，2000年。10 P. de la Harpe. 关于单纯形上的希尔伯特度量。载于 G. A. Niblo 和 M. A. Roller 编，《几何群论》，第97–119页。剑桥大学出版社，1993年。doi:10.1017/CBO9780511661860.009。11 D. DeCoste and B. Schölkopf. 训练不变支持向量机。机器学习，46:161–190，2002年。doi:10.1023/A:1012454411458。12 A. H. Gezalyan and D. M. Mount. 希尔伯特度量中的Voronoi图。载于第39届国际计算几何研讨会论文集，第35:1–35:16页，2023年。doi:10.4230/LIPIcs.SoCG.2023.35。13 S. Huang, N. Cai, P. P. Pacheco, S. Narrandes, Y. Wang, and W. Xu. 支持向量机（SVM）学习在癌症基因组学中的应用。《癌症基因组学与蛋白质组学》，15:41–51，2018年。doi:10.21873/cgp.20063。14 M. Ito, Y. Seo, T. Yamazaki, and M. Yanagida. 关于 Thompson 度量下正定矩阵锥的几何性质。线性代数及其应用，435:2054–2064，2011年。doi:10.1016/j.laa.2011.03.063。15 T. Joachims.《使用支持向量机学习文本分类》。施普林格，2002年。doi:10.1007/978-1-4615-0907-3。

16 N. Karmarkar. 一种新的线性规划多项式时间算法。Combinatorica，4:373–395，1984年。doi:10.1007/BF02579150。17 J. Karwowski and F. Nielsen. 对称正定双锥的希尔伯特几何。载于 NeurIPS 2025 神经表示中的对称性与几何研讨会，2025 年。URL: https://arxiv.org/abs/2508.14369。18 L. G. Khachiyan. 一种线性规划多项式算法。《科学院学报》，244:1093–1096，1979年。URL: https://www.mathnet.ru/eng/dan42319。19 S. Krig. 《计算机视觉度量：综述、分类体系与分析》。Springer Nature，2014年。20 D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá. 复杂网络的双曲几何。Phys. Rev. E，82，2010年。 doi:10.1103/PhysRevE.82.036106。21 W. Lin, V. Duruisseaux, M. Leok, F. Nielsen, M. E. Khan, and M. Schmidt. 简化基于动量的正定子流形优化及其在深度学习中的应用。载于第40届国际机器学习会议论文集，2023年。URL: https://arxiv.org/abs/2302.09738。22 Brian C Lins. 希尔伯特度量非扩张映射的渐近行为与 Denjoy-Wolff定理。博士学位论文，罗格斯大学新不伦瑞克研究生院，2007年。23 J. Matousek and B. Gärtner. 《理解与使用线性规划》。Universitext。施普林格，柏林海德堡，2006年。24 F. Nielsen and K. Sun. 希尔伯特射影几何中的聚类：概率单纯形与相关矩阵椭圆体案例研究。载于 Frank Nielsen 编，《信息的几何结构》，第297–331页。Springer Internat. Pub.，2019年。 doi:10.1007/978-3-030-02520-5_11。25 F. Nielsen and K. Sun. 希尔伯特单纯形几何中的非线性嵌入，2022年。 arXiv:arXiv:2203.11434。26 A. Papadopoulos and M. Troyanov. 《希尔伯特几何手册》，IRMA数学与理论物理讲座第22卷。欧洲数学学会出版社，2014年。 doi:10.4171/147。27 D. Reeb, M. J. Kastoryano, and M. M. Wolf. 量子信息论中的希尔伯特射影度量。J. Math. Physics，52(8)，2011年。doi:10.1063/1.3615729。28 S. J. Russell and P. Norvig. 《人工智能：一种现代方法》。培生，第4版，2021年。29 I. Steinwart and A. Christmann. 《支持向量机》。施普林格，2008年。doi:10.1007/978-0-387-77242-4。30 R. Szeliski. 《计算机视觉：算法与应用》。施普林格，2022年。 doi:10.1007/978-3-030-34372-9。31 V. N. Vapnik and A. Y. Chervonenkis. 关于模式识别学习算法的一类问题。《自动化与遥控》，25(6):937–945，1964年。（俄文）。URL: http://mi.mathnet.ru/at11678。32 C. Vernicos. 关于凸多胞体的希尔伯特几何。载于《希尔伯特几何手册》，IRMA数学与理论物理讲座第22卷，第111–126页。欧洲数学学会出版社，2014年。doi:10.48550/arXiv.1406.0733。