

两种基于自适应投影次梯度的多址MIMO非线性均衡算法设计

贾佳冰, 宋爱民*

大连交通大学理学院, 辽宁 大连

收稿日期: 2025年8月24日; 录用日期: 2025年9月14日; 发布日期: 2025年9月25日

摘要

多输入多输出(MIMO)系统是现代通信技术的核心组成部分, 其中多址MIMO信道的研究具有关键意义。在多址MIMO信道中, 发射端与接收端的多天线配置虽有效提升了传输速率和系统容量, 但信号传输过程中存在的多用户干扰等问题, 导致信道呈现显著的非线性特性, 这给MIMO系统的非线性均衡带来了极大挑战。现有基于自适应投影次梯度法的核自适应多元回归算法, 虽能一定程度上应对上述问题, 却存在稀疏处理能力不足、计算复杂度偏高的缺陷。为此, 本文聚焦多址MIMO系统均衡需求, 以QPSK调制信号为研究载体, 提出两种算法。首先针对传统回归算法的参数耦合问题, 引入二元分类思想, 将信道均衡问题转化为信号类别判别问题, 进而提出多址MIMO核自适应二元分类算法。该算法突破传统回归范式的桎梏, 显著提升计算效率, 验证了回归转分类思路的可行性。进一步针对该算法依赖核矩阵运算的瓶颈, 引入随机傅里叶特征近似传统核计算, 提出多址MIMO基于随机傅里叶特征的自适应二元分类算法。该算法实现非线性问题的线性化转换, 进一步降低计算复杂度。两种算法通过复值信号预测向实部与虚部类别判别的转化, 既保留了核方法对信道非线性的拟合能力, 又规避了传统回归算法的参数耦合问题, 为QPSK调制下的多址MIMO系统信号均衡提供了新的解决方案, 奠定了实验基础。实验结果表明, 所提两种算法相较于现有方法均显著提升计算效率, 且后者进一步缩减字典规模。这一结果验证了算法的有效性与优越性。

关键词

多址MIMO信道, 均衡, 自适应投影次梯度, 核方法, 随机傅里叶特征

Design of Two Multi-Access MIMO Nonlinear Equalization Algorithms Based on Adaptive Projected Subgradient

Jiabing Jia, Aimin Song*

*通讯作者。

文章引用: 贾佳冰, 宋爱民. 两种基于自适应投影次梯度的多址 MIMO 非线性均衡算法设计[J]. 图像与信号处理, 2025, 14(4): 398-411. DOI: 10.12677/jisp.2025.144037

Abstract

Multiple-Input Multiple-Output (MIMO) systems are a core component of modern communication technologies, among which the research on multiple-access MIMO channels holds critical significance. In multiple-access MIMO channels, the multi-antenna configurations at the transmitter and receiver effectively improve the transmission rate and system capacity; however, issues such as multi-user interference during signal transmission lead to significant nonlinear characteristics of the channel, posing great challenges to the nonlinear equalization of MIMO systems. Existing kernel adaptive multivariate regression algorithms based on the adaptive projection subgradient method can address the aforementioned issues to a certain extent, but they suffer from deficiencies, including insufficient sparse processing capability and high computational complexity. To this end, focusing on the equalization requirements of multiple-access MIMO systems and using QPSK-modulated signals as the research carrier, this paper proposes two algorithms. First, to address the parameter coupling issue of traditional regression algorithms, the concept of binary classification is introduced to convert the channel equalization problem into a signal class discrimination problem, thereby proposing a multiple-access MIMO kernel adaptive binary classification algorithm. This algorithm breaks the constraints of the traditional regression paradigm, significantly improves computational efficiency, and verifies the feasibility of the idea of converting regression to classification. Furthermore, to tackle the bottleneck of this algorithm that relies on kernel matrix operations, random Fourier features are introduced to approximate traditional kernel calculations, leading to the proposal of a multiple-access MIMO adaptive binary classification algorithm based on random Fourier features. This algorithm realizes the linearized transformation of nonlinear problems and further reduces computational complexity. By converting complex-valued signal prediction into class discrimination for the real and imaginary parts, both algorithms not only retain the fitting ability of kernel methods for channel nonlinearity but also avoid the parameter coupling problem of traditional regression algorithms, providing a new solution for signal equalization in multiple-access MIMO systems under QPSK modulation and laying the experimental foundation. Experimental results demonstrate that the two proposed algorithms significantly improve computational efficiency compared with existing methods, and the latter further reduces the dictionary size. This result verifies the effectiveness and superiority of the proposed algorithms.

Keywords

Multi-Access MIMO Channel, Equalization, Adaptive Projected Subgradient, Kernel Method, Random Fourier Features

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在现代通信技术体系中,多输入多输出(MIMO)系统凭借在频谱效率与可靠性方面的显著优势,已成为无线通信领域的核心技术之一。多址 MIMO 信道作为复杂通信场景下的关键研究对象,涉及多个发射端借助多天线,同时向配备多天线的接收端传输信号,以此满足持续增长的通信需求。这种配置大幅提

升了数据传输速率与系统容量,但也带来了一系列难题。例如多用户干扰问题,即不同发射端发出的信号在接收端相互干扰,严重妨碍信号的正确接收[1] [2]。

在再生核希尔伯特空间(RKHS)理论持续发展的背景下,研究者专注于探索处理复杂通信问题的有效算法。在文献[3]提出了基于自适应投影次梯度法(APSM)的在线核自适应多元回归算法。APSM 是一种能在在线学习等场景中,通过投影次梯度更新策略,实现参数迭代优化的算法[4]-[6],可用于处理如多址 MIMO 信道信号处理等问题,助力提升系统性能。文献[3]中的算法主要专注于预测连续值信号,通过不断优化回归模型来逼近真实信号值。该过程极为繁琐,且计算复杂度颇高。本文把此问题转化为相对简单的二元分类问题,更着重于信号的类别判断。

此外,在随机傅里叶特征(RFF)的发展脉络下,相关算法针对核方法的计算复杂度瓶颈问题,逐步实现技术突破:在随机傅里叶特征(RFF)的研究脉络中,相关算法已围绕核方法的复杂度与适应性难题持续突破:2019年,Xiong 与 Wang 提出在线随机傅里叶特征共轭梯度(RFFCG)算法,将 RFF 与共轭梯度法结合,通过固定维度映射解决核共轭梯度算法的稀疏化依赖问题,降低了计算与存储复杂度且保持较高滤波精度[7]。2021年,Anand 等人将 RFF 应用于大规模 MIMO 可见光通信,提出随机傅里叶特征-核最小均方(RFF-KLMS)后失真算法,解决了传统 RKHS 方法的字典增长问题,在固定内存预算下实现了与字典式核方法相当的误码率性能并大幅降低计算复杂度[8]。2023年,Gao 等人提出自适应随机傅里叶特征高斯核 LMS (ARFF-GKLMS)算法,通过在线调整 RFF 参数,显著优化了非平稳场景中的收敛率、稳态误差控制与跟踪性能[9]。文献[3]中的算法采用闭球约束与投影映射实现参数稀疏化,结合滑动窗口技术仅留存最近的核函数系数,以此控制模型规模。然而,此类稀疏方法依赖特定参数配置,存在过度剔除重要信息的风险,进而影响模型精度。为此,本文引入 RFF 近似传统核计算:基于 Bochner 定理,从高斯核对应分布采样频率向量,通过三角函数变换构造随机特征向量,使随机特征空间内的内积逼近高斯核计算,摆脱对稀疏化参数的强依赖。

本文聚焦多址 MIMO 系统均衡需求,以 QPSK 调制信号为研究载体,创新体现在两方面:一方面,针对传统回归算法的参数耦合问题,突破核自适应多元回归范式,通过将连续复值信号的回归预测转化为实部与虚部的类别判别,提出多址 MIMO 核自适应二元分类算法,既保留核方法对信道非线性的拟合能力,又显著提升处理效率,验证了回归转分类思路的可行性;另一方面,针对该算法在大规模场景下依赖核矩阵运算的效率瓶颈(数据量增大时,矩阵求逆复杂度随样本量呈指数级增长),引入随机傅里叶特征近似传统核计算,提出多址 MIMO 基于随机傅里叶特征的自适应二元分类算法,实现非线性问题的线性化转换以进一步降低复杂度。实验以误分类率为核心指标验证表明,所提算法较现有方法计算效率显著提升,后者还能压缩字典规模,为多址 MIMO 信道均衡提供了全新解决方案。

本文后续章节安排如下:第二章介绍多址 MIMO 信道模型,后续研究均基于该模型展开;第三章先回顾传统核自适应回归算法,鉴于其计算效率较低,本文从信号类别判断角度切入,提出多址 MIMO 核自适应二元分类算法;第四章为进一步降低计算复杂度,引入随机傅里叶特征近似核函数,提出多址 MIMO 基于随机傅里叶特征的自适应二元分类算法;第五章通过仿真实验,验证所提算法的性能优势;第六章总结全文结论。

2. 多址 MIMO 信道模型

在通信领域,多址 MIMO 信道问题备受关注。本文所研究的多址 MIMO 信道模型包含 P 个用户或发射机(Tx),每个发射机配备 N 根天线,接收机(Rx)配备 M 根天线(符号定义如表 1 所示)。在传输过程中,发射端采用特殊编码策略,如正交空时分组码(OSTBC),充分利用天线分集提升系统性能[10][11]。具体而言,在每个时间瞬间 $n \in \mathbb{N}$,发射端将信息编码为符号块后经信道传输。信息编码向量

$\mathbf{s}_p(n) := [s_{p1}(n), \dots, s_{pK}(n)]^T \in \mathbb{C}^K$, 经编码器 $\mathbf{C}$ 编码后得到 $\mathbf{C}: \mathbb{C}^K \rightarrow \mathbb{C}^{T \times N}: \mathbf{s}_p(n) \mapsto \mathbf{C}(\mathbf{s}_p(n))$ 的符号块, 接收端接收到的信号 $\mathbf{X}(n)$ 由公式

$$\mathbf{X}(n) = \sum_{p=1}^P \mathbf{C}(\mathbf{s}_p(n)) \mathbf{H}_p + \mathbf{V}(n) \in \mathbb{C}^{T \times M}, \forall n \in \mathbb{N} \quad (1)$$

确定, 其中 $\mathbf{H}_p \in \mathbb{C}^{N \times M}$ 为信道矩阵, $\mathbf{V}(n)$ 为噪声。不同方法在编码策略和对信道信息的依赖上存在差异。线性技术通常遵循 OSTBC 编码策略, 且部分或完全依赖信道信息; 而本文所提的非线性核方法采用简单编码策略, $K := T := 1$, 每个天线发送相同符号, 并且不依赖信道信息。

在接收端, 为简化运算并统一符号表示, 对接收信号进行预处理。通过定义映射 $\mathbf{M} := \text{vect } \mathbf{M}$, 其中 vect 表示标准的列堆叠操作, $\mathbf{M}$ 是任意矩阵。基于接收信号 $\mathbf{X}(n), \forall n \in \mathbb{N}$ 得到 $2TM$ 维实向量 $\mathcal{X}_1(n)$ 和 $\mathcal{X}_2(n)$, $\mathcal{X}_1(n)$ 和 $\mathcal{X}_2(n)$ 分别包含 $\mathbf{X}(n)$ 的实部和虚部信息。预处理为

$$\forall n \in \mathbb{N}, \mathcal{X}_1(n) := \begin{bmatrix} \Re(\mathbf{X}(n)) \\ \Im(\mathbf{X}(n)) \end{bmatrix}, \mathcal{X}_2(n) := \begin{bmatrix} \Im(\mathbf{X}(n)) \\ -\Re(\mathbf{X}(n)) \end{bmatrix},$$

其中 $\Re(\cdot)$ 与 $\Im(\cdot)$ 表示复向量的实部和虚部。

Table 1. Notation table
表 1. 符号表

符号	定义
n	时间索引
P	用户/发射机数量
N	每个发射机的天线数量
M	接收机的天线数量
K	编码前的符号数量
T	每个天线发送的符号块数量
$R = K/T$	编码速率
$\mathbf{s}_p(n) \in \mathbb{C}^K$	时刻 n 待编码的 K 个符号组成的向量
$\mathbf{C}$	编码器
$\mathbf{C}(\mathbf{s}_p(n)) \in \mathbb{C}^{T \times N}$	编码后的符号块
$\{\mathbf{H}_p\}_{p=1}^P \in \mathbb{C}^{N \times M}$	信道矩阵
L	感兴趣用户/发射机数量
$\mathbf{M}$	矩阵 $\mathbf{M}$ 列堆叠得到的向量

3. 多址 MIMO 问题的核自适应二元分类算法设计

鉴于传统核自适应多元回归算法在多址 MIMO 场景中存在计算效率偏低的局限, 本章将先对该算法进行简要回顾, 进而展开核自适应二元分类算法的设计与推导。

3.1. 回顾传统核自适应多元回归算法

文献[3]提出了一个适用于在线监督多元回归任务的通用框架, 该框架构建于无限维 RKHS 之上, 并应用于资源受限且无信道先验信息的多址 MIMO 信道均衡场景。此框架将回归问题建模在无限维 RKHS

中, 可覆盖线性模型及多种非线性多元回归模型, 允许使用任何凸的、连续的、不一定可微的函数作为损失函数, 只需其次梯度具有解析形式。考虑到后续仿真实验需引用该算法, 为降低文本冗余, 本文将命名为“多址 MIMO 问题的核自适应多元回归算法(Kernel Adaptive Multivariate Regression Algorithm, 简记为 KAMR)”。引入笛卡尔积空间 $\mathcal{H} := \mathcal{H}_1 \times \mathcal{H}_2 \times \cdots \times \mathcal{H}_L$, 其元素 $\mathbf{f} := [f_1, f_2, \dots, f_L]^T$, 且 $f_i \in \mathcal{H}_i$ 。通过正定矩阵 $\mathbf{P} = [p_{ij}] \in \mathbb{R}^{L \times L}$ 定义内积为

$$\langle \langle \mathbf{f}_1, \mathbf{f}_2 \rangle \rangle := \sum_{i,j=1}^L p_{ij} \langle f_{1i}, f_{2j} \rangle, \mathbf{f}_1, \mathbf{f}_2 \in \mathcal{H}, \quad (2)$$

相应的, $\mathcal{H}$ 中的诱导范数定义为 $\|\cdot\| := \sqrt{\langle \langle \cdot, \cdot \rangle \rangle}$ 。该算法中接收端定义为映射

$$\mathbf{f} : \mathbb{R}^{2TM} \rightarrow \mathbb{R}^{KL} : \mathbb{R}^{2M} \rightarrow \mathbb{R}^L (K := T := 1) : \mathbf{x} \mapsto \mathbf{f}(\mathbf{x}) := [f_1(\mathbf{x}), \dots, f_L(\mathbf{x})]^T, \quad (3)$$

其中每个分量 f_i 属于 RKHS。目标信号为 $\mathbf{s}(n)$ 。接收端处理后的信号为 $\mathbf{f}(\mathcal{X}_1(n)) + i\mathbf{f}(\mathcal{X}_2(n))$ 。目标是找到 $\mathbf{f}$ 使得 $\mathbf{f}(\mathcal{X}_1(n)) - \Re(\mathbf{s}(n)), \mathbf{f}(\mathcal{X}_2(n)) - \Im(\mathbf{s}(n))$ 足够小。通过最小化损失函数, 可以使接收端处理后的信号尽可能接近原始发送信号, 从而提高信号恢复的准确性。算法的大致过程为: 初始化阶段选择非负参数 $\varepsilon \geq 0$ 和窗口大小 q , 构建 ε -不敏感损失函数 $l_\varepsilon(\xi) := \max\{0, l(\xi) - \varepsilon\} \forall \xi \in \mathbb{R}^L$, 仅对超出容忍度的误差施加惩罚, 从而增强模型鲁棒性。算法通过滑动窗口 $\mathcal{J}_n := \max\{0, n - q + 1\}, n$ 动态选取最近 q 个时间步的训练数据, 并对每个数据点 $j \in \mathcal{J}_n$ 计算损失函数 $\mathcal{L}_{\varepsilon,j}(\mathbf{f}_n)$ 关于 $\mathbf{f}_n$ 的次梯度 $\mathcal{L}'_{\varepsilon,j}(\mathbf{f}_n)$ (形式依 ℓ_1 -范数)。定义活跃索引集 $\mathcal{I}_n := \{j \in \mathcal{J}_n : \mathcal{L}'_{\varepsilon,j}(\mathbf{f}_n) \neq \mathbf{0}\}$, 为活跃损失函数分配归一化权重 $\{\omega_i^{(n)}\}_{i \in \mathcal{I}_n} \subset (0, 1]$, 按以下公式更新多元回归估计值

$$\mathbf{f}_{n+1} := \mathbf{f}_n - \mu_n \sum_{i \in \mathcal{I}_n} \omega_i^{(n)} \frac{\mathcal{L}_{\varepsilon,i}(\mathbf{f}_n)}{\|\mathcal{L}'_{\varepsilon,i}(\mathbf{f}_n)\|^2} \mathcal{L}'_{\varepsilon,i}(\mathbf{f}_n). \quad (4)$$

该算法通过不断迭代, 逐步调整多回归器, 使其在处理在线多元回归问题时能够更好地适应数据变化, 渐近最小化损失函数序列。

针对在线核方法中内存和计算负载随时间增长的问题, 文献[3]采用了闭球投影与滑动缓冲区稀疏策略。在本文中我们将这个带有稀疏策略的算法称作“基于闭球投影与滑动缓冲区稀疏的多址 MIMO 核自适应多元回归算法(Kernel Adaptive Multivariate Regression Algorithm Based on Sparsity via Closed Ball Projection and Sliding Buffer, 简记为 KAMR-CBP-SB)”。在执行稀疏化操作时, 就可能错误地将对模型精度起关键支撑作用的重要信息过度剔除, 破坏数据内在关联, 使模型对输入的拟合与预测出现偏差, 降低整体精度。与此同时, 传统核自适应多元回归算法在仿真实验中, 暴露出计算效率低下的显著缺陷, 当数据规模扩大时, 难以满足高效计算的实际需求。鉴于此, 需探索更具效能的解决方案。

3.2. 多址 MIMO 问题的核自适应二元分类算法

传统核自适应多元回归应用于多址 MIMO 问题时, 聚焦于连续值信号预测, 需通过不断优化回归模型, 经历复杂的迭代计算来逼近真实信号值。但在实际场景中, 像区分不同用户发送的信号、识别不同调制方式的信号这类任务, 更核心的需求是对信号类别进行判断。相较于连续值预测的繁琐过程, 直接将其转化为二元分类问题处理, 可大幅简化流程并显著提升处理效率。

假设存在 L 个 RKHS $\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_L$, 笛卡尔积空间为 $\mathcal{H} := \mathcal{H}_1 \times \mathcal{H}_2 \times \cdots \times \mathcal{H}_L$ 。在这样的空间架构下, 本文所研究的分类器可以巧妙地建模为笛卡尔积空间 $\mathcal{H}$ 中的元素, 即

$$\mathbf{f} := [f_1, f_2, \dots, f_L]^T, f_1 \in \mathcal{H}_1, \dots, f_L \in \mathcal{H}_L,$$

与 $\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_L$ 相关联的正定核分别是 $\kappa_1(\cdot, \cdot), \kappa_2(\cdot, \cdot), \dots, \kappa_L(\cdot, \cdot)$, 每个正定核都对应着一个特定的核宽度 $\sigma_1, \sigma_2, \dots, \sigma_L$, 且 $(\cdot)^t$ 表示向量/矩阵转置。受文献[3]启发, 结合本文研究中无限维再生核希尔伯特空间的内积构造需求, 现将内积与范数定义为如下形式:

$$\langle\langle f_1, f_2 \rangle\rangle := \sum_{i=1}^L \langle f_{1i}, f_{2i} \rangle_{\mathcal{H}_i}. \quad (5)$$

相应的, $\mathcal{H}$ 中的诱导范数定义为 $\|f\| := \sqrt{\langle\langle f, f \rangle\rangle}$ 。

在笛卡尔积空间的基础上, 我们进一步探索其中的半空间结构。半空间的研究对于解决分类问题中的边界划分和优化求解具有重要意义。在实际的分类任务中, 为了衡量分类结果与真实标签之间的差异, 我们定义损失函数为:

$$l_{x,y,\rho}(f): \mathcal{H} \rightarrow \mathbb{R}: f \mapsto \left(\rho - \sum_{i=1}^L y_i f_i(x) \right)^+ \quad (6)$$

其中 $(\cdot)^+$ 表示取非负值。定义 $\mathcal{H}$ 中的半空间为:

$$\Pi_{x,y,\rho}^+ = \left\{ f \in \mathcal{H} : \langle\langle f, a_{x,y} \rangle\rangle_{\mathcal{H}} \geq \rho \right\}, \quad (7)$$

记 $\kappa(x, \cdot) = [\kappa_1(x, \cdot), \dots, \kappa_L(x, \cdot)]^t$, 其中 $a_{x,y} := y \odot \kappa(x, \cdot) \in \mathcal{H}$ 表示法向量, 这里的符号 $\odot$ 是哈达玛积, 即对于两个相同尺寸的矩阵/向量 A 与 B , $A \odot B$ 表示在 A 和 B 相同位置上的元素相乘的结果。

值得一提的是, 半空间与损失函数之间存在着紧密的内在联系, 这种联系在后续的算法设计和优化过程中起着关键作用, 它能够帮助我们更好地理解分类问题的本质和优化方向。下面介绍两个命题。

命题 1: 设笛卡尔积空间 $\mathcal{H} := \mathcal{H}_1 \times \mathcal{H}_2 \times \dots \times \mathcal{H}_L$ 中, 定义其内积为(5)式, 诱导范数为 $\|f\| := \sqrt{\langle\langle f, f \rangle\rangle_{\mathcal{H}}}$ 。

给定半空间 $\Pi_{x_n, y_n, \rho}^+ = \left\{ f_n \in \mathcal{H} : \langle\langle f_n, a_{x_n, y_n} \rangle\rangle_{\mathcal{H}} \geq \rho \right\}$, 其中 $a_{x_n, y_n} := y_n \odot \kappa(x_n, \cdot) \in \mathcal{H}$ 为法向量。那么对于任意 $f_n \in \mathcal{H}$, 其在半空间 $\Pi_{x_n, y_n, \rho}^+$ 上的投影为:

$$P_{\Pi_{x_n, y_n, \rho}^+}(f_n) = f_n + \frac{\left(\rho - \langle\langle f_n, a_{x_n, y_n} \rangle\rangle_{\mathcal{H}} \right)^+}{\|a_{x_n, y_n}\|^2} a_{x_n, y_n}. \quad (8)$$

由于

$$\begin{aligned} \|a_{x,y}\|^2 &= \|y \odot \kappa(x, \cdot)\|^2 = \langle\langle y \odot \kappa(x, \cdot), y \odot \kappa(x, \cdot) \rangle\rangle_{\mathcal{H}} \\ &= \sum_{i=1}^L \langle y_i \kappa_i(x, \cdot), y_i \kappa_i(x, \cdot) \rangle = \sum_{i=1}^L \kappa_i(x, x) = L. \end{aligned}$$

则有

$$P_{\Pi_{x_n, y_n, \rho}^+}(f_n) = f_n + \frac{1}{L} \left(\rho - \langle\langle f_n, a_{x_n, y_n} \rangle\rangle_{\mathcal{H}} \right)^+ y_n \odot \kappa(x_n, \cdot). \quad (9)$$

证明: 设投影后的向量为 $f \in \Pi_{x_n, y_n, \rho}^+$ 我们的目标是最小化 f_n 与 f 的距离, 即最小化 $\|f_n - f\|^2$, 同时满足约束条件 $\langle\langle f, a_{x_n, y_n} \rangle\rangle_{\mathcal{H}} \geq \rho$ 。根据 $\mathcal{H}$ 中诱导范数的定义 $\|f\| := \sqrt{\langle\langle f, f \rangle\rangle}$, 则优化问题可表示为:

$$\min_{f \in \Pi_{x_n, y_n, \rho}^+} \langle\langle f_n - f, f_n - f \rangle\rangle_{\mathcal{H}} \quad \text{s.t.} \quad \langle\langle f, a_{x_n, y_n} \rangle\rangle_{\mathcal{H}} - \rho \geq 0.$$

若 f_n 已经满足约束 $\langle\langle f, a_{x_n, y_n} \rangle\rangle_{\mathcal{H}} - \rho \geq 0$, 那么 f_n 本身就是该优化问题的最优解, 也就是说,

$P_{\Pi_{x_n, y_n, \rho}^+}(\mathbf{f}_n) = \mathbf{f}_n$; 若 $\mathbf{f}_n$ 不满足约束, 为了处理约束条件, 引入非负拉格朗日乘子 $\lambda \geq 0$, 构建拉格朗日函数:

$$\mathcal{L}(\mathbf{f}, \lambda) = \langle \langle \mathbf{f}_n - \mathbf{f}, \mathbf{f}_n - \mathbf{f} \rangle \rangle_{\mathcal{H}} - \lambda \left(\langle \langle \mathbf{f}, \mathbf{a}_{x_n, y_n} \rangle \rangle_{\mathcal{H}} - \rho \right).$$

根据内积定义将 $\langle \langle \mathbf{f}_n - \mathbf{f}, \mathbf{f}_n - \mathbf{f} \rangle \rangle_{\mathcal{H}}$ 展开得:

$$\langle \langle \mathbf{f}_n - \mathbf{f}, \mathbf{f}_n - \mathbf{f} \rangle \rangle_{\mathcal{H}} = \sum_{i=1}^L \langle \langle f_{n,i} - f_i, f_{n,i} - f_i \rangle \rangle_{\mathcal{H}_i} = \sum_{i=1}^L \left(\langle \langle f_{n,i}, f_{n,i} \rangle \rangle_{\mathcal{H}_i} - 2 \langle \langle f_{n,i}, f_i \rangle \rangle_{\mathcal{H}_i} + \langle \langle f_i, f_i \rangle \rangle_{\mathcal{H}_i} \right).$$

所以拉格朗日函数可进一步写为:

$$\mathcal{L}(\mathbf{f}, \lambda) = \sum_{i=1}^L \left(\langle \langle f_{n,i}, f_{n,i} \rangle \rangle_{\mathcal{H}_i} - 2 \langle \langle f_{n,i}, f_i \rangle \rangle_{\mathcal{H}_i} + \langle \langle f_i, f_i \rangle \rangle_{\mathcal{H}_i} \right) - \lambda \left(\sum_{i=1}^L \langle \langle f_i, \mathbf{a}_i \rangle \rangle_{\mathcal{H}_i} - \rho \right).$$

其中 $\mathbf{a}_i$ 表示 $\mathbf{a}_{x_n, y_n}$ 的第 i 个分量。对 $\mathbf{f}$ 的每个分量 $f_i (i \in 1, 2, \dots, L)$ 求偏导数且令偏导数为 0, 有

$$\frac{\partial \mathcal{L}(\mathbf{f}, \lambda)}{\partial f_i} = -2f_{n,i} + 2f_i - \lambda \mathbf{a}_i = 0.$$

对于每个 $i \in 1, 2, \dots, L$, 都有

$$f_i = f_{n,i} + \frac{\lambda}{2} \mathbf{a}_i.$$

则向量 $\mathbf{f}$ 可表示为:

$$\mathbf{f} = \mathbf{f}_n + \frac{\lambda}{2} \mathbf{a}_{x_n, y_n}.$$

对 λ 求偏导并令其为 0, 可以得到

$$\langle \langle \mathbf{f}, \mathbf{a}_{x_n, y_n} \rangle \rangle_{\mathcal{H}} \geq \rho.$$

将 $\mathbf{f}$ 代入, 得到

$$\left\langle \left\langle \mathbf{f}_n + \frac{\lambda}{2} \mathbf{a}_{x_n, y_n}, \mathbf{a}_{x_n, y_n} \right\rangle \right\rangle_{\mathcal{H}} \geq \rho.$$

结合内积与范数定义, 得到

$$\langle \langle \mathbf{f}_n, \mathbf{a}_{x_n, y_n} \rangle \rangle_{\mathcal{H}} + \frac{\lambda}{2} \|\mathbf{a}_{x_n, y_n}\|^2 \geq \rho.$$

解得

$$\lambda = \frac{2 \left(\rho - \langle \langle \mathbf{f}_n, \mathbf{a}_{x_n, y_n} \rangle \rangle_{\mathcal{H}} \right)^+}{\|\mathbf{a}_{x_n, y_n}\|^2}.$$

将 λ 代入到 $\mathbf{f}$ 表达式, 最终得到投影公式:

$$P_{\Pi_{x_n, y_n, \rho}^+}(\mathbf{f}_n) = \mathbf{f}_n + \frac{\left(\rho - \langle \langle \mathbf{f}_n, \mathbf{a}_{x_n, y_n} \rangle \rangle_{\mathcal{H}} \right)^+}{\|\mathbf{a}_{x_n, y_n}\|^2} \mathbf{a}_{x_n, y_n}.$$

□

命题 2: 在再生核希尔伯特空间的分类问题中, 闭半空间 $\Pi_{\mathbf{x}_n, \mathbf{y}_n, \rho}^+$ 与软边际损失函数 $l_{\mathbf{x}_n, \mathbf{y}_n, \rho}(\mathbf{f}_n)$ 的所有极小值点的集合是一致的。即:

$$\arg \min_{\mathbf{f}_n \in \mathcal{H}} l_{\mathbf{x}_n, \mathbf{y}_n, \rho}(\mathbf{f}_n) = \Pi_{\mathbf{x}_n, \mathbf{y}_n, \rho}^+. \quad (10)$$

证明: 由于 $\Pi_{\mathbf{x}_n, \mathbf{y}_n, \rho}^+$ 是一个闭半空间, 其相关的度量投影映射 $P_{\Pi_{\mathbf{x}_n, \mathbf{y}_n, \rho}^+}(\mathbf{f}_n)$ 为:

$$P_{\Pi_{\mathbf{x}_n, \mathbf{y}_n, \rho}^+}(\mathbf{f}_n) = \mathbf{f}_n + \frac{1}{L} \left(\rho - \left\langle \mathbf{f}_n, \mathbf{a}_{\mathbf{x}_n, \mathbf{y}_n} \right\rangle_{\mathcal{H}} \right)^+ \mathbf{y}_n \odot \boldsymbol{\kappa}(\mathbf{x}_n, \cdot).$$

由距离函数的定义

$$d(\mathbf{f}, S) := \inf \left\{ \left\| \mathbf{f} - \hat{\mathbf{f}} \right\| : \hat{\mathbf{f}} \in S \right\},$$

假设 $\mathcal{H}$ 为再生核希尔伯特空间, 其中 S 是 $\mathcal{H}$ 的子集。对于任意 $\mathbf{f} \in \mathcal{H}$, 有

$$d(\mathbf{f}_n, \Pi_{\mathbf{x}_n, \mathbf{y}_n, \rho}^+) = \left\| \mathbf{f}_n - P_{\Pi_{\mathbf{x}_n, \mathbf{y}_n, \rho}^+}(\mathbf{f}_n) \right\|,$$

因此

$$\begin{aligned} d(\mathbf{f}_n, \Pi_{\mathbf{x}_n, \mathbf{y}_n, \rho}^+) &= \left\| \mathbf{f}_n - P_{\Pi_{\mathbf{x}_n, \mathbf{y}_n, \rho}^+}(\mathbf{f}_n) \right\| = \left\| \frac{1}{L} \left(\rho - \left\langle \mathbf{f}_n, \mathbf{a}_{\mathbf{x}_n, \mathbf{y}_n} \right\rangle_{\mathcal{H}} \right)^+ \mathbf{y}_n \odot \boldsymbol{\kappa}(\mathbf{x}_n, \cdot) \right\| \\ &= \frac{1}{L} \left(\rho - \left\langle \mathbf{f}_n, \mathbf{a}_{\mathbf{x}_n, \mathbf{y}_n} \right\rangle_{\mathcal{H}} \right)^+ \left\langle \mathbf{y}_n \odot \boldsymbol{\kappa}(\mathbf{x}_n, \cdot), \mathbf{y}_n \odot \boldsymbol{\kappa}(\mathbf{x}_n, \cdot) \right\rangle_{\mathcal{H}} \\ &= \frac{1}{L} \left(\rho - \left\langle \mathbf{f}_n, \mathbf{a}_{\mathbf{x}_n, \mathbf{y}_n} \right\rangle_{\mathcal{H}} \right)^+ L = l_{\mathbf{x}_n, \mathbf{y}_n, \rho}(\mathbf{f}_n). \end{aligned}$$

这表明最小化损失函数 $l_{\mathbf{x}_n, \mathbf{y}_n, \rho}(\mathbf{f}_n)$ 的问题等价于最小化 $d(\cdot, \Pi_{\mathbf{x}_n, \mathbf{y}_n, \rho}^+)$ 。然而, $d(\cdot, \Pi_{\mathbf{x}_n, \mathbf{y}_n, \rho}^+)$ 的所有极小值点的集合是 $\Pi_{\mathbf{x}_n, \mathbf{y}_n, \rho}^+$ 。□

基于上述命题, 本节提出多址 MIMO 问题的核自适应二元分类算法(Kernel Adaptive Binary Classification Algorithm, 简记为 KABC), 其依托自适应投影次梯度法(APSM)实现。APSM 支持在每个时刻对多个数据进行并行处理以提高收敛速度。具体而言, 在时刻 n , 并非仅处理单个数据对 $(\mathbf{x}_n, \mathbf{y}_n)$, 而是处理集合 $\{(\mathbf{x}_n, \mathbf{y}_n)\}_{j \in \mathcal{J}_n}$, 其中指标集 $\mathcal{J}_n \subset \overline{n-q+1, n}$ ($q \in \mathbb{Z}_{>0}$ 用于界定先前时刻接收的数据对范围)。为简化研究, 假设再生核希尔伯特空间满足 $\mathcal{H}_1 = \mathcal{H}_2 = \dots = \mathcal{H}_L$, 且各空间均采用实高斯核函数 $\kappa_{\mathbb{R}}$, 其表达式为

$$\kappa_{\mathbb{R}}(\mathbf{x}_i, \mathbf{x}_j) = \exp \left(\frac{-\left\| \mathbf{x}_i - \mathbf{x}_j \right\|^2}{2\sigma^2} \right), \text{ 记由 } L \text{ 个该核函数组成的向量为 } \boldsymbol{\kappa}(\cdot, \cdot) = [\kappa_{\mathbb{R}}(\cdot, \cdot), \kappa_{\mathbb{R}}(\cdot, \cdot), \dots, \kappa_{\mathbb{R}}(\cdot, \cdot)]_L^T. \text{ 针对}$$

每个数据对 $(\mathbf{x}_n, \mathbf{y}_n)$ 及指标 $j \in \mathcal{J}_n$, 定义半空间 $\Pi_{\mathbf{x}_j, \mathbf{y}_j, \rho}^+$, 为简化符号, 统一记为 $\Pi_{j, n}^+$, 其形式为:

$$\Pi_{j, n}^+ := \left\{ \mathbf{f}_n \in \mathcal{H} : \left\langle \mathbf{f}_n, \mathbf{a}_{\mathbf{x}_j, \mathbf{y}_j} \right\rangle_{\mathcal{H}} \geq \rho \right\}, \quad (11)$$

每个半空间 $\Pi_{j, n}^+$ 都与一个度量投影映射 $P_{\Pi_{j, n}^+}$ 相关联, 表达式为:

$$P_{\Pi_{j, n}^+}(\mathbf{f}_n) = \mathbf{f}_n + \frac{\left(\rho - \left\langle \mathbf{f}_n, \mathbf{a}_{\mathbf{x}_j, \mathbf{y}_j} \right\rangle_{\mathcal{H}} \right)^+}{\left\| \mathbf{a}_{\mathbf{x}_j, \mathbf{y}_j} \right\|^2} \mathbf{y}_j \odot \boldsymbol{\kappa}(\mathbf{x}_j, \cdot). \quad (12)$$

在每个迭代轮次 n , 一旦接收到 $(\mathbf{x}_n, \mathbf{y}_n)$, 将来自先前迭代的可用估计投影到子空间 $\Pi_{j,n}^+$ 上, 并通过适当的权重组合这些投影。为简化权重设计假设输出的 L 维共享同一权重。为每个 $j \in \mathcal{J}_n$ 关联一个权重 $w_j^{(n)} \geq 0$, 满足 $\sum_{j \in \mathcal{J}_n} w_j^{(n)} = 1$ 。考虑初始分类器点 $\mathbf{f}_0 \in \mathcal{H}$, 并通过以下方式生成 $\mathcal{H}$ 中的分类器序列:

$$\mathbf{f}_{n+1} := \mathbf{f}_n + \mu_n \left(\sum_{j \in \mathcal{J}_n} w_j^{(n)} P_{\Pi_{j,n}^+}(\mathbf{f}_n) - \mathbf{f}_n \right), \forall n \in \mathbb{Z}_{\geq 0}, \quad (13)$$

其中外推系数 $\mu_n \in [0, 2\mathcal{M}_n]$, 且

$$\mathcal{M}_n := \begin{cases} \frac{\sum_{j \in \mathcal{J}_n} w_j^{(n)} \|P_{\Pi_{j,n}^+}(\mathbf{f}_n) - \mathbf{f}_n\|^2}{\left\| \sum_{j \in \mathcal{J}_n} w_j^{(n)} P_{\Pi_{j,n}^+}(\mathbf{f}_n) - \mathbf{f}_n \right\|^2}, & \text{如果 } \mathbf{f}_n \notin \bigcap_{j \in \mathcal{J}_n} \Pi_{j,n}^+, \\ 1, & \text{否则.} \end{cases} \quad (14)$$

这里 $\mathcal{M}_n \geq 1, \forall n \in \mathbb{Z}_{\geq 0}$, 所以外推系数 μ_n 可以取大于或等于 2 的值。对于 $\forall j \in \mathcal{J}_n, \forall n \in \mathbb{Z}_{\geq 0}$, 定义

$$\boldsymbol{\beta}_j^{(n)} := \frac{\left(\rho - \left\langle \mathbf{f}_n, \mathbf{a}_{\mathbf{x}_j, \mathbf{y}_j} \right\rangle_{\mathcal{H}} \right)^+ w_j^{(n)}}{\left\| \mathbf{a}_{\mathbf{x}_n, \mathbf{y}_n} \right\|^2} \mathbf{y}_j = \frac{\left(\rho - \left\langle \mathbf{f}_n, \mathbf{a}_{\mathbf{x}_j, \mathbf{y}_j} \right\rangle_{\mathcal{H}} \right)^+ w_j^{(n)}}{L} \mathbf{y}_j, \quad (15)$$

其中 $\mathbf{y}_j = [y_{1j}^{(n)}, y_{2j}^{(n)}, \dots, y_{Lj}^{(n)}]^t$ 。将式(11)与式(12)代入到式(13)中可以得到:

$$\begin{aligned} \mathbf{f}_{n+1} &:= \mathbf{f}_n + \mu_n \left(\sum_{j \in \mathcal{J}_n} w_j^{(n)} P_{\Pi_{j,n}^+}(\mathbf{f}_n) - \mathbf{f}_n \right), \forall n \in \mathbb{Z}_{\geq 0} \\ &= \mathbf{f}_n + \mu_n \left(\sum_{j \in \mathcal{J}_n} w_j^{(n)} \left[\mathbf{f}_n + \frac{\left(\rho - \left\langle \mathbf{f}_n, \mathbf{a}_{\mathbf{x}_j, \mathbf{y}_j} \right\rangle_{\mathcal{H}} \right)^+}{\left\| \mathbf{a}_{\mathbf{x}_n, \mathbf{y}_n} \right\|^2} \mathbf{y}_j \odot \boldsymbol{\kappa}(\mathbf{x}_j, \cdot) \right] - \mathbf{f}_n \right) \\ &= \mathbf{f}_n + \mu_n \left(\sum_{j \in \mathcal{J}_n} w_j^{(n)} \mathbf{f}_n + \sum_{j \in \mathcal{J}_n} w_j^{(n)} \frac{\left(\rho - \left\langle \mathbf{f}_n, \mathbf{a}_{\mathbf{x}_j, \mathbf{y}_j} \right\rangle_{\mathcal{H}} \right)^+}{\left\| \mathbf{a}_{\mathbf{x}_n, \mathbf{y}_n} \right\|^2} \mathbf{y}_j \odot \boldsymbol{\kappa}(\mathbf{x}_j, \cdot) - \mathbf{f}_n \right) \\ &= \mathbf{f}_n + \mu_n \left(\sum_{j \in \mathcal{J}_n} w_j^{(n)} \mathbf{f}_n + \sum_{j \in \mathcal{J}_n} \frac{\left(\rho - \left\langle \mathbf{f}_n, \mathbf{a}_{\mathbf{x}_j, \mathbf{y}_j} \right\rangle_{\mathcal{H}} \right)^+ w_j^{(n)}}{\left\| \mathbf{a}_{\mathbf{x}_n, \mathbf{y}_n} \right\|^2} \mathbf{y}_j \odot \boldsymbol{\kappa}(\mathbf{x}_j, \cdot) - \mathbf{f}_n \right) \\ &= \mathbf{f}_n + \mu_n \left(\mathbf{f}_n + \sum_{j \in \mathcal{J}_n} \frac{\left(\rho - \left\langle \mathbf{f}_n, \mathbf{a}_{\mathbf{x}_j, \mathbf{y}_j} \right\rangle_{\mathcal{H}} \right)^+ w_j^{(n)}}{\left\| \mathbf{a}_{\mathbf{x}_n, \mathbf{y}_n} \right\|^2} \mathbf{y}_j \odot \boldsymbol{\kappa}(\mathbf{x}_j, \cdot) - \mathbf{f}_n \right) \\ &= \mathbf{f}_n + \mu_n \left(\sum_{j \in \mathcal{J}_n} \boldsymbol{\beta}_j^{(n)} \odot \boldsymbol{\kappa}(\mathbf{x}_j, \cdot) \right) \\ &= \mathbf{f}_n + \mu_n \sum_{j \in \mathcal{J}_n} \boldsymbol{\beta}_j^{(n)} \odot \boldsymbol{\kappa}(\mathbf{x}_j, \cdot). \end{aligned}$$

那么算法的迭代更新公式可以等价地写为:

$$\mathbf{f}_{n+1} = \mathbf{f}_n + \mu_n \sum_{j \in \mathcal{J}_n} \boldsymbol{\beta}_j^{(n)} \odot \boldsymbol{\kappa}(\mathbf{x}_j, \cdot). \quad (16)$$

经过适当的代数运算, 参数 $\mathcal{M}_n$ 取以下形式:

$$\mathcal{M}_n := \begin{cases} \frac{\sum_{j \in \mathcal{J}_n} w_j^{(n)} \left\| \left(\rho - \left\langle \mathbf{f}_n, \mathbf{a}_{\mathbf{x}_j, \mathbf{y}_j} \right\rangle_{\mathcal{H}} \right)^+ \right\|^2}{\sum_{i, j \in \mathcal{J}_n} \boldsymbol{\beta}_i^{(n)} \boldsymbol{\beta}_j^{(n)} \kappa(\mathbf{x}_i, \mathbf{x}_j)}, & \text{如果 } \mathbf{f}_n \notin \bigcap_{j \in \mathcal{J}_n} \Pi_{j,n}^+, \\ 1, & \text{否则.} \end{cases}$$

4. 基于随机傅里叶特征的多址 MIMO 非线性算法设计

4.1. 随机傅里叶特征理论基础

为降低核方法的计算复杂度, 考虑引入随机傅里叶特征(Random Fourier Features, RFF)。其核心思想是利用有限维随机特征向量, 对高斯核的无限维映射进行近似, 从而将核方法的高复杂度计算转化为低维空间的线性运算。对于高斯核 $\kappa(\mathbf{x}, \mathbf{x}') = \exp\left(\frac{-\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2}\right)$ 这一平移不变核, 根据 Bochner 定理(正定核可分解为随机频率向量的三角函数期望), 其可展开为

$$\kappa(\mathbf{x}, \mathbf{x}') \approx E_{\boldsymbol{\omega}, b} [\cos(\boldsymbol{\omega}' \mathbf{x} + b) \cos(\boldsymbol{\omega}' \mathbf{x}' + b)],$$

其中随机频率向量 $\boldsymbol{\omega} \sim \mathcal{N}(0, \sigma^{-2} \mathbf{I})$ ($\mathcal{N}(0, \sigma^{-2} \mathbf{I})$ 表示均值为 0、协方差为 $\sigma^{-2} \mathbf{I}$ 的多元正态分布), 相位 $b \sim \text{Uniform}([0, 2\pi])$ (均匀分布于 $[0, 2\pi]$)。通过随机采样 D 组独立样本 $\{(\boldsymbol{\omega}_i, b_i)\}_{i=1}^D$, 构造特征映射:

$$\mathbf{z}_{\Omega}(\mathbf{x}) = \sqrt{\frac{2}{D}} [\cos(\boldsymbol{\omega}_1' \mathbf{x} + b_1), \cos(\boldsymbol{\omega}_2' \mathbf{x} + b_2), \dots, \cos(\boldsymbol{\omega}_D' \mathbf{x} + b_D)]^T.$$

符号 $\Omega = \{\boldsymbol{\omega}_1, b_1, \dots, \boldsymbol{\omega}_D, b_D\}$, 即随机特征参数集合(含随机频率 $\boldsymbol{\omega}_k$ 和相位 b_k , $k = 1, 2, \dots, D$)。最终高斯核的内积运算可近似为低维特征映射的内积:

$$\kappa(\mathbf{x}, \mathbf{x}') \approx \mathbf{z}_{\Omega}(\mathbf{x})^T \mathbf{z}_{\Omega}(\mathbf{x}').$$

4.2. 基于随机傅里叶特征的多址 MIMO 算法

本节提出多址 MIMO 基于随机傅里叶特征的自适应二元分类算法(Adaptive Binary Classification Algorithm Based on Random Fourier Features, 简记为 RFF-ABC)。首先通过随机傅里叶特征, 将无限维函数映射到有限维参数空间。构造随机特征映射 $\mathbf{z}_{\Omega}(\mathbf{x})$ 为:

$$\mathbf{z}_{\Omega}(\mathbf{x}) = \sqrt{\frac{2}{D}} \begin{bmatrix} \cos(\boldsymbol{\omega}_1' \mathbf{x} + b_1) \\ \cos(\boldsymbol{\omega}_2' \mathbf{x} + b_2) \\ \vdots \\ \cos(\boldsymbol{\omega}_D' \mathbf{x} + b_D) \end{bmatrix}, \quad (17)$$

其中, $\boldsymbol{\omega}_k \sim \mathcal{N}(0, \sigma^{-2} \mathbf{I})$, $k = 1, 2, \dots, D$, 服从均值为 0、协方差 $\sigma^{-2} \mathbf{I}$ 的高斯分布, σ 是核宽度参数;

$b_k \sim \text{Uniform}([0, 2\pi])$ (随机相位, 均匀分布在 $[0, 2\pi]$); D 是随机特征维度, 通过该映射, $z_\Omega: \mathbb{R}^d \rightarrow \mathbb{R}^D$, 把输入数据映射到 D 维欧氏空间。

假设分类器为 $\mathbf{f} := [f_1, f_2, \dots, f_L]^T \in \mathbb{R}^L$, 利用 RFF 的有限维近似, 假设 L 维空间的随机特征映射完全相同, 每个空间函数可表示为

$$f_k = W_k^T z_\Omega(\cdot),$$

其中 $W_k \in \mathbb{R}^D$ 为系数列向量, $z_\Omega(\cdot) \in \mathbb{R}^D$ 为特征映射列向量。进一步, 将系数列向量 $\{W_k\}_{k=1}^L$ 按列拼接为系数矩阵:

$$\mathbf{W} := [W_1, W_2, \dots, W_L] \in \mathbb{R}^{D \times L},$$

此时, 分类器可紧凑表示为如下形式:

$$\mathbf{f}(\cdot) = \mathbf{W}^T z_\Omega(\cdot). \quad (18)$$

由于采用随机傅里叶特征近似核方法, 将无限维再生核希尔伯特空间(RKHS)的非线性问题转化为欧氏空间的线性化框架, 因此基于随机傅里叶特征的有限维映射, 将内积定义为: 对任意 $W_1, W_2 \in \mathbb{R}^{D \times L}$, 设其第 k 列分别为 $W_{1,k}, W_{2,k} \in \mathbb{R}^D$, 则内积为:

$$\langle W_1, W_2 \rangle := \sum_{k=1}^L W_{1,k}^T W_{2,k}. \quad (19)$$

对应范数为:

$$\|\mathbf{W}\| := \sqrt{\sum_{k=1}^L W_k^T W_k}. \quad (20)$$

假设分类标签为 $\mathbf{y} = [y_1, y_2, \dots, y_L]^T \in \mathbb{R}^L$, 满足 $|y_k| = 1$, 特征映射为 $z_\Omega(\mathbf{x}) \in \mathbb{R}^D$, 分类器系数矩阵 $\mathbf{W} \in \mathbb{R}^{D \times L}$ (其第 k 列为 $W_k \in \mathbb{R}^D$) 需满足的半空间约束为:

$$\Pi_{\mathbf{x}, \mathbf{y}, \rho}^+ = \left\{ \mathbf{W} \in \mathbb{R}^{D \times L} : \sum_{k=1}^L W_k^T \tilde{a}_k \geq \rho \right\}, \quad (21)$$

其中, 法向量矩阵为: $\tilde{\mathbf{a}}_{\mathbf{x}, \mathbf{y}} := [y_1 z_\Omega(\mathbf{x}), y_2 z_\Omega(\mathbf{x}), \dots, y_L z_\Omega(\mathbf{x})] \in \mathbb{R}^{D \times L}$, 其第 k 列 $\tilde{a}_k = y_k z_\Omega(\mathbf{x}) \in \mathbb{R}^D$ 对应第 k 维的法向量分量。对于任意系数矩阵 $\mathbf{W} \in \mathbb{R}^{D \times L}$ 其在该半空间上的投影可表示为:

$$P_{\Pi_{\mathbf{x}, \mathbf{y}, \rho}^+}(\mathbf{W}_n) = \mathbf{W}_n + \frac{\left(\rho - \sum_{k=1}^L W_k^T \tilde{a}_k \right)^+}{\|\tilde{\mathbf{a}}_{\mathbf{x}, \mathbf{y}}\|^2} \tilde{\mathbf{a}}_{\mathbf{x}, \mathbf{y}}. \quad (22)$$

式中分子的 $(\cdot)^+$ 表示取非负值, 分母 $\|\tilde{\mathbf{a}}_{\mathbf{x}, \mathbf{y}}\|^2$ 是法向量矩阵的范数平方, 利用 $|y_k| = 1$, 即 $y_k^2 = 1$, 该范数平方可简化推导为:

$$\begin{aligned} \|\tilde{\mathbf{a}}_{\mathbf{x}, \mathbf{y}}\|^2 &= \sum_{k=1}^L \tilde{a}_k^T \tilde{a}_k = \sum_{k=1}^L (y_k z_\Omega(\mathbf{x}))^T (y_k z_\Omega(\mathbf{x})) \\ &= \sum_{k=1}^L y_k^2 z_\Omega(\mathbf{x})^T z_\Omega(\mathbf{x}) = \sum_{k=1}^L z_\Omega(\mathbf{x})^T z_\Omega(\mathbf{x}) = L z_\Omega(\mathbf{x})^T z_\Omega(\mathbf{x}). \end{aligned}$$

在线学习场景中, 对于每个时刻 $n \in \mathbb{Z}_{\geq 0}$ 都关联着一对 $(\mathbf{x}_n, \mathbf{y}_n)$ 。每个时刻并行处理来自先前 q 长度时间窗口的数据集 $\{(\mathbf{x}_n, \mathbf{y}_n)\}_{j \in \mathcal{J}_n}$, 其中指标集 $\mathcal{J}_n \subset \overline{n-q+1, n}$, 为简化权重设计假设输出的 L 维共享同一权重

重。为每个 $j \in \mathcal{J}_n$ 关联一个权重 $w_j^{(n)} \geq 0$, 满足 $\sum_{j \in \mathcal{J}_n} w_j^{(n)} = 1$ 。定义投影系数:

$$\beta_j^{(n)} := \frac{\left(\rho - \sum_{k=1}^L W_k^t \tilde{a}_k\right)^+ w_j^{(n)}}{\|\tilde{a}_{x_j, y_j}\|} y_j = \frac{\left(\rho - \sum_{k=1}^L W_k^t \tilde{a}_k\right)^+ w_j^{(n)}}{Lz_\Omega(\mathbf{x}_j)^t z_\Omega(\mathbf{x}_j)} y_j, \quad (23)$$

其中标签 $\mathbf{y}_j = [y_{1j}^{(n)}, y_{2j}^{(n)}, \dots, y_{Lj}^{(n)}]^t$, 法向量矩阵为 $\tilde{a}_{x_j, y_j} = [y_{1j}^{(n)} z_\Omega(\mathbf{x}_j), y_{2j}^{(n)} z_\Omega(\mathbf{x}_j), \dots, y_{Lj}^{(n)} z_\Omega(\mathbf{x}_j)] \in \mathbb{R}^{D \times L}$, 其第 k 列 $\tilde{a}_k^{(j)} = y_{kj}^{(n)} z_\Omega(\mathbf{x}_j) \in \mathbb{R}^D$ 对应第 k 维的法向量分量, 范数 $\|\tilde{a}_{x_j, y_j}\| = Lz_\Omega(\mathbf{x}_j)^t z_\Omega(\mathbf{x}_j)$ 。系数矩阵更新遵循如下规则:

$$\mathbf{W}_{n+1} = \mathbf{W}_n + \mu_n \sum_{j \in \mathcal{J}_n} \left(\beta_j^{(n)} z_\Omega(\mathbf{x}_j)^t \right)^t, \quad (24)$$

式中外推系数 $\mu_n \in [0, 2\mathcal{M}_n]$, 其自适应调整机制为:

$$\mathcal{M}_n := \begin{cases} \frac{\sum_{j \in \mathcal{J}_n} w_j^{(n)} \|P_{\Pi_{j,n}^+}(\mathbf{W}_n) - \mathbf{W}_n\|^2}{\left\| \sum_{j \in \mathcal{J}_n} w_j^{(n)} P_{\Pi_{j,n}^+}(\mathbf{W}_n) - \mathbf{W}_n \right\|^2}, & \text{如果 } \mathbf{W}_n \notin \bigcap_{j \in \mathcal{J}_n} \Pi_{j,n}^+, \\ 1, & \text{否则.} \end{cases}$$

该机制通过投影误差动态调控 $\mathcal{M}_n$, 保证更新步长合理性。

5. 仿真实验

5.1. 实验参数

数据与验证设置: 加载 MIMO 系统数据集, 设定训练数据量为 3000 个, 测试数据量为 50 个; 实验重复 10 次, 以此确保结果稳定; 采用固定间隔验证策略, 每 50 个训练样本进行一次验证, 验证点数量为训练数据量与验证间隔的比值。

MIMO 系统参数配置: 用户/发射机数量 P 为 4、每个发射机天线数量 N 为 2、接收机天线数量 M 为 2、感兴趣用户/信号数量 L 为 2、发射机每次编码前待编码符号块复符号数量 K 为 1、每个天线发送的符号块数 T 为 1。每个发射端的源符号 $\mathbf{s}_p(n)$ 的各分量等概率地取 $\pm 1 \pm i$ 这些复数值。信道矩阵 $\mathbf{H}_p$ 的每个分量都是复值随机变量, 其实部和虚部均服从均值为 0、方差为 1 的正态分布。并且同一信道矩阵内的不同分量相互独立, 不同信道矩阵的分量之间也相互独立。

5.2. 实验结果

仿真实验结果如下图所示。

基于图 1 的仿真结果可知: 核自适应二元分类算法(KABC)相较于核自适应多元回归算法(KAMR), 计算复杂度降低显著, 相同任务下运行耗时远低于 KAMR; 且从误分类率指标来看, KABC 的分类性能未因复杂度优化而衰减, 实现了计算复杂度与分类性能的协同保持。需说明的是, KABC 因未引入稀疏化策略, 在字典规模上存在固有局限。进一步分析随机傅里叶特征(RFF)近似的 RFF-ABC: 其相较于 KABC 进一步缩减字典规模与计算复杂度, 且在同等字典规模约束下, 耗时低于采用闭球投影-滑动窗稀疏策略的 KAMR-CBP-SB; 但局限性在于, 相较于未引入稀疏机制的 KABC, 其误分类率呈小幅上升趋势。对于 RFF-ABC 的核心参数(维数 D)的选取, 若需在 RFF-ABC 现有复杂度基础上进一步降低计算复杂度,

需减小维数 D , 但 D 的进一步缩减会导致误分类率升高; 反之, 若要提升误分类性能, 则需增大 D , 而 D 增大又会推高计算复杂度, 二者存在权衡制约关系。

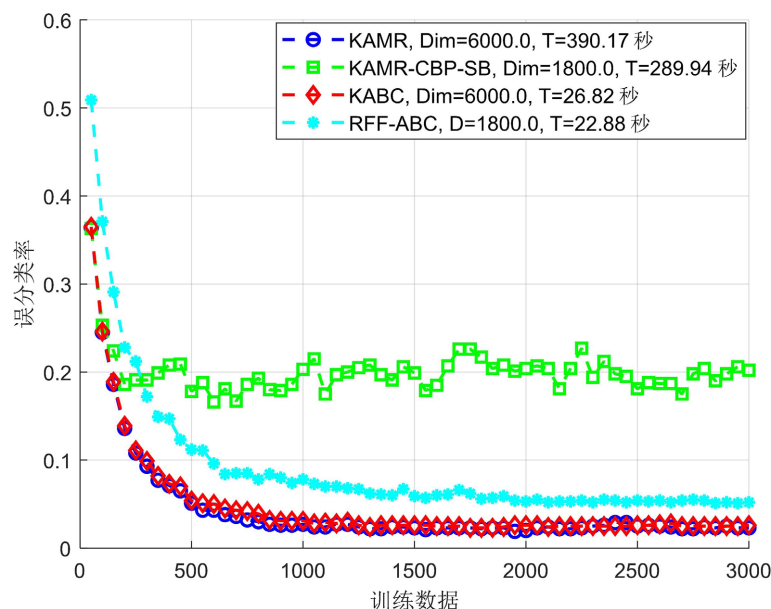


Figure 1. Simulation experiment result diagrams of this paper
图 1. 本文仿真实验结果图

6. 结论

本文围绕多址 MIMO 系统的信号处理问题, 以 QPSK 调制信号为研究对象, 借助自适应投影次梯度法开展研究: 一方面, 将传统聚焦于连续值预测的复值信号处理, 转化为实部与虚部的类别判别, 既保留核方法对信道非线性的拟合能力, 又规避传统回归算法的参数耦合问题, 摒弃复杂的连续值预测流程、提升计算效率, 验证了回归转分类的可行性; 另一方面, 引入随机傅里叶特征近似传统核计算, 通过构造随机特征映射将高斯核非线性问题转化为线性问题, 在降低计算复杂度、减小字典规模的同时实现较低误分类率, 优化模型性能。仿真结果表明, 两种方法均显著提升计算效率, 其中随机傅里叶特征方案可进一步压缩字典规模, 为 QPSK 调制下的多址 MIMO 系统信道均衡提供了新的解决方案。

参考文献

- [1] Shahbazpanahi, S., Beheshti, M., Gershman, A.B., Gharavi-Alkhansari, M. and Wong, K.M. (2004) Minimum Variance Linear Receivers for Multiaccess MIMO Wireless Systems with Space-Time Block Coding. *IEEE Transactions on Signal Processing*, **52**, 3306-3313. <https://doi.org/10.1109/tsp.2004.837442>
- [2] Cavalcante, R.L.G. and Yamada, I. (2008) Multiaccess Interference Suppression in Orthogonal Space-Time Block Coded MIMO Systems by Adaptive Projected Subgradient Method. *IEEE Transactions on Signal Processing*, **56**, 1028-1042. <https://doi.org/10.1109/tsp.2007.907896>
- [3] Slavakis, K., Bouboulis, P. and Theodoridis, S. (2012) Adaptive Multiregression in Reproducing Kernel Hilbert Spaces: The Multiaccess MIMO Channel Case. *IEEE Transactions on Neural Networks and Learning Systems*, **23**, 260-276. <https://doi.org/10.1109/tnnls.2011.2178321>
- [4] Slavakis, K., Theodoridis, S. and Yamada, I. (2008) Online Kernel-Based Classification Using Adaptive Projection Algorithms. *IEEE Transactions on Signal Processing*, **56**, 2781-2796. <https://doi.org/10.1109/tsp.2008.917376>
- [5] Slavakis, K., Yamada, I., Ogura, N. and Yukawa, M. (2004) Adaptive Projected Subgradient Method and Set Theoretic Adaptive Filtering with Multiple Convex Constraints. *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, 2004, Pacific Grove, 7-10 November 2004, 960-964.

- <https://doi.org/10.1109/acssc.2004.1399281>
- [6] Fink, J., Cavalcante, R.L.G. and Stańczak, S. (2023) Superiorized Adaptive Projected Subgradient Method with Application to MIMO Detection. *IEEE Transactions on Signal Processing*, **71**, 1350-1362. <https://doi.org/10.1109/tsp.2023.3263255>
- [7] Xiong, K. and Wang, S. (2019) The Online Random Fourier Features Conjugate Gradient Algorithm. *IEEE Signal Processing Letters*, **26**, 740-744. <https://doi.org/10.1109/lsp.2019.2907480>
- [8] Anand, P.K., Jain, S., Mitra, R. and Bhatia, V. (2021) Random Fourier Features Based Post-Distortion for Massive-Mimo Visible Light Communication. 2020 *International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*, Sharjah, 16-18 March 2021, 1-6. <https://doi.org/10.1109/iccspa49915.2021.9385752>
- [9] Gao, W., Chen, J., Richard, C., Shi, W. and Zhang, Q. (2023) Adaptive Random Fourier Features Kernel LMS. 2023 *IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, Zhengzhou, 14-17 November 2023, 1-6. <https://doi.org/10.1109/icspcc59353.2023.10400238>
- [10] Alamouti, S.M. (1998) A Simple Transmit Diversity Technique for Wireless Communications. *IEEE Journal on Selected Areas in Communications*, **16**, 1451-1458. <https://doi.org/10.1109/49.730453>
- [11] Tarokh, V., Jafarkhani, H. and Calderbank, A.R. (1999) Space-Time Block Codes from Orthogonal Designs. *IEEE Transactions on Information Theory*, **45**, 1456-1467. <https://doi.org/10.1109/18.771146>