

重庆邮电大学研究生综述报告

报告题目	基于核逻辑回归的遥感图像分类方法
学 号	S170231053
姓 名	唐建烺
学位层次	硕士研究生
学位类别	工程硕士
学科专业	计算机技术
研究方向	智能信息处理
研究生导	雷大江 副教授
报告日期	2018 年 11 月 24 日
报告分数	
导师签名	
学院审核	

基于核逻辑回归的遥感图像分类方法

唐建烺

重庆邮电大学计算机科学与技术学院 重庆 400065

E-mail: donjy@foxmail.com

联系电话: 18883336267

摘要: 逻辑回归作为一个经典的分类算法,其局限性在于只能作用于线性可分的数据集。而对于线性不可分的数据集,则可以通过内核方法将数据映射到更高维的空间,希望在这个更高维空间中变得更容易分离或结构化。逻辑回归经过核化扩充之后的分类器称之为核 Logistic 回归 (Kernel logistic regression, KLR),这样就变成了可用于非线性特征数据分类的分类器。本文利用传统的单核学习将所有的样本数据通过一个核函数映射到一个高维特征空间上去,使其在这个高维空间上线性可分。遥感图像分类在军事安全和自然灾害监测上有着广泛的应用,但是遥感图像分类存在两个主要挑战:一是数据存在非线性可分的情况;二是数据的空间特征和光谱特征通常是异构的,基于单个特征空间的单核分类器很难满足分类需求。因此本文又将多核学习和逻辑回归算法相结合,对遥感图像中的空间特征和光谱特征采用不同的核函数进行映射,使数据中不同类型的特征在不同的高维空间中能够更好的表达。但是,随着数据的不断增大,核方法的使用越来越受限制。在大规模数据的情况下,存储和计算核矩阵的开销都非常的大。本文通过低秩近似核矩阵再结合快速对偶算法不仅加速了 KLR 算法的求解过程,还去掉了数据中的冗余信息提高了分类准确率。

关键字: 核 Logistic 回归, 多核学习, 遥感图像分类, 低秩近似, 对偶算法

Remote Sensing Image Classification based on Kernel Logistic Regression

Abstract: Logistic regression as a classic classification algorithm, has limitations in that it can only be applied to linearly separable datasets. For linearly indivisible datasets, we can map data to higher-dimensional space through kernel trick. It becomes easier to separate or structure in this higher dimensional space. The logistic regression is called kernel logistic regression (KLR) after the kernel expansion, it becomes a classifier that can be used for nonlinear feature data classification. In this paper, all sample data are mapped to a high-dimensional feature space through a kernel function by using traditional single-kernel learning, so that it can be linearly separable in this high-dimensional space. The remote sensing image classification has a wide range of applications in military safety and natural disaster monitoring. However, there are two main challenges in remote sensing image classification: one is the existence of non-linear separability of data; the other is that spatial and spectral features of data are

usually heterogeneous. The single-kernel classifiers based on a single feature space are difficult to meet classification requirements. Therefore, this paper combines multi-kernel learning with logistic regression algorithm to map spatial and spectral features of remote sensing images using different kernel functions, so that different types of features in data can be better expressed in different high-dimensional space. With the scale of data continues to expand, and the application of kernel trick is increasingly restricted. In the case of large-scale data, the cost of storage and computing kernel matrix is very large. In this paper, the low-rank approximate kernel matrix combined with the fast dual algorithm not only accelerates the solution of the KLR algorithm but also removes the redundant information in the data to improve the classification accuracy.

Keywords: multiple kernel logistic regression, multiple kernel learning, remote sensing image classification, low-rank approximate, dual algorithm

1. 引言

1.1 研究的背景及意义

逻辑回归(Logistic Regression, LR)算法是统计分析、机器学习和数据挖掘领域的一个经典分类算法,主要针对二分类问题提出,用于估计事件发生的可能性,输出等价于模型预测某个数据点属于正类的概率估计。它是一种线性的分类算法,具有求解速度快、预测结果可解释性强的特点,它为探索性解释数据提供了一个有用的概率模型。但是,一般的 LR 只能用于线性分类,对于非线性的特征数据分类就十分的困难。Logistic 回归经过核化扩充之后的分类器称之为核 Logistic 回归(Kernel logistic regression, KLR),这样就变成了可用于非线性特征数据分类的分类器^[1]。通过将原数据的样本映射到一个高维或无穷维的特征空间,使得样本在这个特征空间内线性可分^[2, 3]。虽然此过程的映射规则通常是不可知的,但是可以用核函数(Kernel function)去代替特征空间上的内积运算,因此并不需要知道此特征空间里特征向量的具体表达形式^[4, 5]。

核方法^[6]在近年来机器学习领域十分流行,尤其是基于统计学习理论的支持向量机^[7]的成功,由此产生了多种基于核函数的方法^[8],并且被广泛的应用于模式识别、文本分类、信号处理等多个领域^[9]。相比于支持向量机, KLR 的目标函数不是风险最小函数,而是最大似然值。因此 KLR 的解会产生类别分类的后验概率值,同时 KLR 也是一个凸优化问题^[10],局部最优解一定是全局最优解。而对于一个凸优化问题,可以运用梯度法或牛顿迭代法进行求解^[11]。但是该两种方法在每次迭代过程中都需要对一个 $n \times n$ 的核矩阵进行求逆等运算,其中 n 表示样本个数,通常 n 的数量达到几千个时,计算时间代价就变得非常的大,甚至难以

接受，其时间复杂度为 $O(n^3)$ 。在支持向量机的凸二次规划问题求解中，Platt 等人提出了序列最小最优化算法（Sequential minimal optimization, SMO）^[12]，基于此算法思想的启发，Keerthi 等人又在 2005 年给出了一种用于计算 KLR 的快速对偶算法，该算法不需要将整个核矩阵带入迭代步骤里进行计算，每一次迭代只优化序列里的两个值，避免了对核矩阵进行求逆等操作，因此迭代的计算代价十分的小^[13]。单核学习是将所有的样本特征通过一个核函数映射到一个高维空间上去。在使用核函数的时候，需要我们根据经验或者试验来选择哪种核函数，以及选取它的参数。如果分类器只采用基于单个特征空间的一个核函数将所有数据映射到一个高维空间，组成一个庞大的高维特征，显然这样对于某些具有异构特征信息的数据是不合理的，而且特征数据通常不是单分布的。而通过多核学习，将不同类型的异构特征采用不同的核函数或者不同参数的核组成多个核函数，然后训练每个核的权重，选出最佳核函数组合再扩充到分类器中，这样就能提高核函数的灵活性，发挥各个基本核的不同特征映射能力，从而提高分类精度。

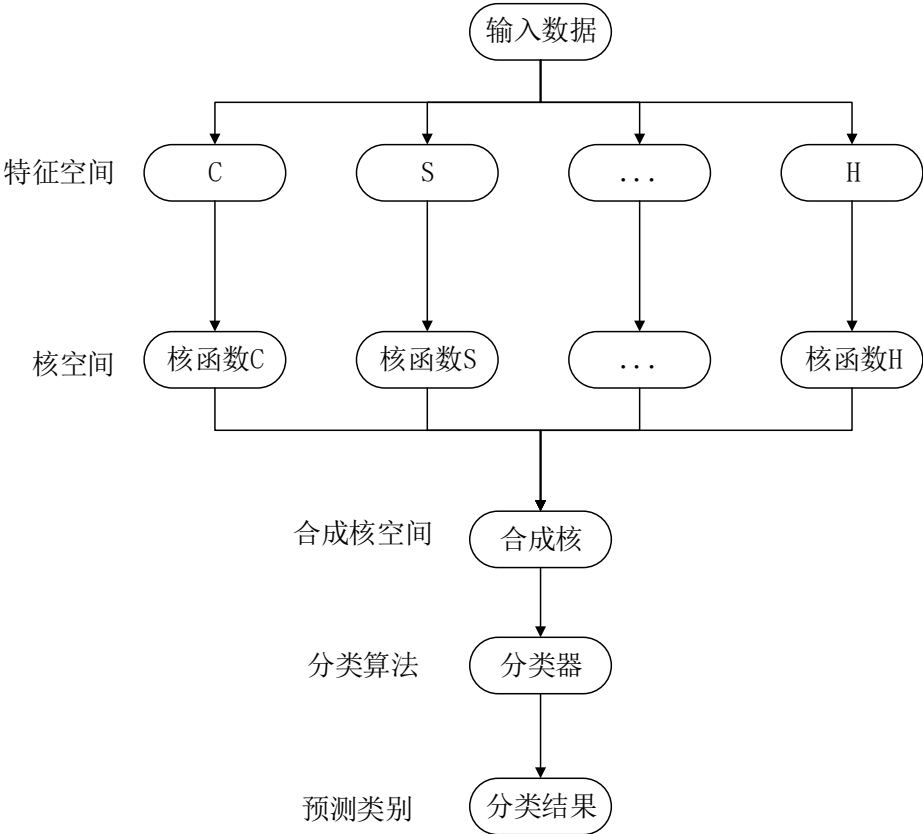


图 1 多核学习算法基本流程图

随着互联网技术的不断发展，数据规模的不断扩大，核方法的应用越来越受限制。其中一个关键的问题是核矩阵通常都是稠密矩阵，其存储和计算代价都非常的高，存储稠密矩阵需要 $O(n^2)$ 的空间，而计算这样的矩阵则需要 $O(n^2m)$ 的代价，这里 n 和 m 分别代表了样本的个数和维度^[14]。最常用的解决方法是用有限的

内存来计算得到一个近似的核矩阵^[15, 16]，这种方法不仅解决了内存问题，也加快了核矩阵的计算。目前已有多种核近似方法被提出^[17, 18]，其中 Nystrom 方法的应用最为广泛，也是之后很多求解近似矩阵算法的基础^[19]。Nystrom 方法的主要思想是通过降低秩来得到原始矩阵 K 的近似矩阵 \tilde{K} ，相当于从 K 中随机取 m 行 m 列，则 $\tilde{K} = K_{n,m} K_{m,m}^{-1} K_{m,n}$ ，其中 $K_{n,m}$ 是原始矩阵 K 中的 $n \times m$ 块矩阵，而 $m \ll n$ 。本文将 KLR 的快速对偶算法、梯度下降算法、牛顿迭代法^[20]和近似求解核矩阵的 Nystrom 方法相结合，提出了不同的基于近似核矩阵的快速求解 KLR 的算法。

遥感图像的识别和分类目前被广泛的应用于资源调查、自然灾害检测、天气气象预告和国防安全上。遥感图像分类是指根据不同地物的电磁辐射信息特征判别该地物的类别及其特征。遥感图像数据主要是通过遥感卫星或者航天遥感技术来采集的，遥感技术在军事方面有很多的应用，其主要包括军事侦察、导弹预警、武器制导、军事测绘等，是目前最为有效、最为安全，同时又是最可靠的侦察手段，对国家安全、社会稳定和军事战略布局起着至关重要的作用。而遥感图像内容的标注是比较困难的，通常需要人工标注。目前神经网络在图像识别和分类中有着广泛的应用，但是对于神经网络的训练需要大量的样本，而获取的遥感图像数据已标注的训练样本集较小，在少量样本的情况下，神经网络不仅资源消耗大而且训练效果不明显。多核学习在小样本的学习中，通过对不同的异构特征加入不同的核函数将其映射到多个高维特征空间中，并引入惩罚因子产生一定的软边界，可以在可观的时间上获取相当高的分类准确率。

1.2 国内外研究现状

核方法最早是应用的支持向量机(Support Vector Machine, SVM)上,通过引入核函数来解决非线性分类问题。在 SVM 获得巨大的成功中，核方法扮演了非常重要的角色，进一步的，核方法被应用到了其他传统的学习算法中去，可以很方便的把线性算法转化为非线性算法。例如：核主成分分析(Kernel PCA)^[21]、核 Fisher 判别分析(Kernel FDA)^[22]、核聚类分析等等。这些基于核的学习算法都简称为核方法，只采用一个基本核的学习方法称之为单核学习。相比于单个核函数，多核模型具有更高的灵活性。多核学习方法根据不同的分类标准有不同的分类方式，按照多核函数的构造方法和特点的不同，可以将多核学习方法大致分成三大类别：合成核方法、多尺度核方法、无限核方法。而多核学习理论早期大体上分为：基于 Boosting 的多核组合模型学习方法，基于半定规划(Semi-definite programming, SDP) 的多核学习方法，基于二次约束型二次规划(Quadratically constrained quadratic program, QCQP)的学习方法，基于半无限线性规划(Semi-definite linear program, SILP)的学习方法和基于超核(Hyper kernels)的学习方法等

[23]。而近年来出现的简单多核学习(Simple MKL)方法[24]和以及基于中心对齐的学习核算法(Algorithms for Learning Kernels Based on Centered Alignment)[25]都是比较流行的多核学习方法。

Logistic regression 算法是一种广义的线性回归分析模型,为了解决非线性分类问题,于是将核方法引入到了该分类器中,便出现了核逻辑回归(Kernel logistic regression, KLR)。但是由于引入了核矩阵,求解 KLR 非常的缓慢。受到 Platt 的序列最小最优算法(Sequential Minimal Optimization, SMO)的启发,Keerthi 基于对偶理论[26]提出了求解 KLR 的快速对偶算法[13],避免了将核矩阵带入到迭代过程中去运算,从而极大的减少了计算开销。但是随着数据量的不断增大,核矩阵的计算也会影响整个算法的开销,并且将整个数据样本都通过核函数映射到高维空间上去,会带入一些冗余信息。因此,矩阵的近似求解方法被广泛的应用在了求解核矩阵上,如:Nystrom 方法, Fastfood 方法[27]等。通过这些方法可以加速核矩阵的求解,利用低秩近似可以去掉原始数据中的一些冗余信息。

随着卫星遥感图像和航空遥感图片分辨率的不断提高,人们可以从遥感图像中获得更多的有用的数据和信息。遥感图像是按一定比例尺,客观真实地记录和反映地表物体的电磁辐射的强弱信息,是遥感探测所获得的遥感信息资料的一种表现形式。所以,遥感技术应用的核心问题是根据地物辐射电磁辐射强弱在遥感图像上表现的特征,判读识别地面物体的类属及其分布特征。遥感图像特征取决于遥感探测通道、地物光谱特征、大气传播特征及传感器的响应特征等因素。只要了解这些因素对遥感图像特征的影响,则可按图像特征判读地面物体的属性及其分布范围,实现遥感图像的分类识别。目前图像识别和分类方法通常都是采用神经网络。但是在实际应用中,遥感图像的数据的采集和标定是较困难的,获取的已标定的训练样本集较小,卷积神经网络需要动辄百万级的训练样本,小样本训练效果往往不能满足需求,如何从小量的数据中生成良好的神经网络或者对小样本异构特征图像进行识别和分类将是未来的研究方向。此外迁移学习也可以解决遥感数据样本少的问题。而多核学习则可以将不同类型的异构特征进行不同的映射处理,充分利用了遥感图像空间结构特征,能在有效的时间内得到可观的分类准确率[28]。

2. 综述主题

2.1 研究目标

(1) 拟提出一种快速核逻辑回归算法。针对逻辑回归算法不能作用于非线性数据集的问题,将单核学习方法与其结合,并采用对偶算法、梯度下降算法和

牛顿迭代法分别进行求解。

(2) 拟提出一种多核多元逻辑回归算法。将核逻辑回归算法扩展成多分类器，使其能适用于多类别的分类任务，并增加稀疏项，拟从效率和精度上进行改进。

(3) 拟提出一种基于核逻辑回归算法的遥感图像分类方法。由于遥感图像的多特征特性，通过设计多核多元逻辑回归算法去提升分类效果。

(4) 拟设计一个遥感图像分类的原型系统。通过提前训练好模型，然后将上传的遥感图像进行分类，识别图像中地物的类别并展示出来。

2.2 研究内容

2.2.1 快速核逻辑回归算法

逻辑回归是一种有监督的常用分类算法，但是不能处理非线性特征的数据。为了解决非线性数据分类问题，通过核方法将逻辑回归进行核扩充后将原始样本映射到高维空间上去再进行分类。在求解核逻辑回归算法过程中，为了避免将核矩阵带入到迭代运算过程中，于是采用 SMO 优化算法去求解，减少算法的时间复杂度。引入核方法后，在求解核逻辑回归(KLR)时，随着样本规模不断扩大，核方法的应用越来越受限制。核矩阵通常都是稠密矩阵，其存储和计算代价都非常的高，为了解决这个问题，采用 Nystrom 方法去低秩近似核矩阵，一方面不仅能减少核矩阵的运算开销，另一方面还能通过低秩减少数据中的冗余信息，提高分类器的精度。

2.2.2 多核多元逻辑回归算法

单核学习的核函数只有一个，其结构单一，对于某些具有不同类型特征的异构数据，如果都统一采用一个核函数，映射到同一个特征空间上去，显然是不合理的。多核学习则是给定一些基本的核函数，对于每一个基核函数又可以指定不同的参数，最后通过它们的加权线性组合来得到最终的复核函数。由于融合了多个核函数，使得不同类型的特征数据映射到不同的高维空间中能更好的表达。如何得到多核学习的线性组合权重，最经典的就是 simpleMKL，被应用在了计算机各个领域。为了使多核学习应用地更广，应对各种特征组合、各种样本量级，后来提出了 GMKL，最优化方法用的是 PGD (Projected Gradient Descend)。而基于中心对齐的学习核算法(Algorithms for Learning Kernels Based on Centered Alignment)^[29]是比较流行的多核学习方法，通过学习能得到一组核函数的权重系数，由此可以组合成新的核函数。

一般的逻辑回归只能作用于二分类，为了适用于多类别的多分类任务，需要将其扩充成多元逻辑回归。单核学习对于某些具有不同类型的数据不够适用，于是将多核学习引入到多元逻辑回归中，并增加稀疏项，提高算法的效率和精度。

2.2.3 基于核逻辑回归算法的遥感图像分类方法

对于遥感图像数据，由于数据的维度很高，数据类型也通常是异构的，容易产生大量的冗余信息。如果在选取核函数的时候，只选用一种核函数会导致遥感图像丰富的信息难以被有效利用。为了解决这个问题，本文采用多核学习方法，经过多个核函数将不同类型的特征数据映射到多个高维特征空间上，最终使得数据在新的组合空间上能够更加准确、合理的表达。于是提出了基于多核多元稀疏逻辑回归算法的遥感图像分类方法。

2.2.4 实现遥感图像分类的原型系统

拟采用 B/S 架构实现基于多核多元稀疏逻辑回归算法的遥感图像分类原型系统,通过提前预训练好模型,然后对新上传的遥感图像进行像素点级别的分类,最后将图像中的地物以不同的颜色标注并可视化。

3. 目前存在的问题

(1) 对于核逻辑回归算法的求解,通常都会引入一个稠密的核矩阵,其存储和计算代价都非常的高,存储稠密矩阵需要 $O(n^2)$ 的空间,而计算这样的矩阵则需要 $O(n^2m)$ 的代价,这里 n 和 m 分别代表了样本的个数和维度。如果直接将全部样本加载到内存中计算所需的核矩阵,通常会引起内存溢出,很难处理大规模的数据集。

(2) 在求解核逻辑回归算法的时候,如果采用梯度下降或者牛顿迭代法这样的迭代求解算法,在其迭代过程中核矩阵都会参与迭代运算,并且还可能涉及到对核矩阵求逆等操作。当数据达到一定规模的时候,核矩阵的大小是和样本个数成正比的,对大型的核矩阵进行矩阵迭代运算开销十分的大,有时候甚至难以接受。

(3) 由于采用了多核学习,这里就会涉及到合成核的问题。把具有不同类型特征的多个核函数进行组合来得到新的多核函数。多核组合方法主要是将基本的核函数直接进行求和或者加权求和,这样就有可能丢失掉原始数据的某些特征信息。

(4) 由于核矩阵的计算开销都十分的大,虽然采用了低秩近似的方法去加速求解核矩阵,但是在选取近似核矩阵的秩的时候需要格外谨慎。秩的选取可能会导致近似核矩阵是奇异矩阵,无法直接进行求逆操作,最终可能导致带入到分类器中后算法不收敛的情况。如果秩选取得太大,则反而会增加求解近似核矩阵的开销,导致分类器最后的训练时间增加,得不偿失。采用低秩近似方法,虽然减少了数据的冗余信息,但是也可能丢失了一些原始数据的相关信息,因此在部分数据集上可能会导致分类准确率有所下降,这样的加速分类器就会以牺牲准确率为代价。

(5) 对于遥感图像的分类问题,由于遥感图像中不仅包含光谱特征还有空间信息特征,如果只利用其中的光谱特征对其像素点进行分类,精度通常难以达到需求。因此还应当充分的利用其空间特征。

4. 进一步的工作

根据上述存在的问题,可以从以下方面进行改进工作。

(1) 针对逻辑回归算法不能作用于非线性数据集的问题, 将单核学习方法与其结合。采用对偶算法、梯度下降算法和牛顿迭代法进行求解 KLR, 并比较以上几种算法的性能。

(2) 在低秩近似核矩阵方法上, 采用了 Nystrom 方法。秩的选取上给定了一个最大阈值的秩, 然后在这个范围内去选取不超过这个给定超参数秩大小的近似核矩阵的秩。Nystrom 方法的主要思想是通过降低秩来得到原始矩阵 K 的近似矩阵 \tilde{K} , 相当于从 K 中随机取 m 行 m 列, 去生成需要的近似核矩阵。

(3) 针对遥感图像中不同类型特征的异构数据, 采用多核学习的方法与 KLR 算法相结合, 用不同的核函数去映射遥感图像中的光谱特征和空间特征。在多核组合方法上, 采用了基于中心对齐的 align 方法^[25]去学习内核的线性组合, 从而产生组合的内核矩阵, 增加了核方法的灵活性, 提高了泛化能力。

参考文献

- [1] Ngufor C, Wojtusiak J. Extreme logistic regression[J]. Advances in Data Analysis and Classification, 2016, 10(1):27-52.
- [2] Mackenzie C A, Trafalis TB, Barker K. A Bayesian beta kernel model for binary classification and online learning problems[J]. Statistical Analysis and Data Mining, 2014, 7(6):434-449.
- [3] Yang L, Qian Y. A sparse logistic regression framework by difference of convex functions programming[J]. Applied Intelligence, 2016, 45(2):1-14.
- [4] AVRON H, CLARKSON K L, WOODRUFF D P J S J O M A, et al. Faster Kernel Ridge Regression Using Sketching and Preconditioning [J]. SIAM Journal on Matrix Analysis and Applications, 2017, 38(4): 1116-38. Avron H, Clarkson K L, Woodruff D P. Faster kernel ridge regression using sketching and preconditioning[J]. SIAM Journal on Matrix Analysis and Applications, 2017, 38(4): 1116-1138.
- [5] Koppel A, Warnell G, Stump E, et al. Parsimonious online learning with kernels via sparse projections in function space[C]. Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on. IEEE, 2017: 4671-4675.
- [6] Kang Z, Peng C, Cheng Q. Kernel-driven similarity learning[J]. Neurocomputing, 2017, 267: 210-219.
- [7] Vapnik V. The nature of statistical learning theory[M]. Springer science & business media, 2013.

- [8] Zhou J, Duan B, Huang J, et al. Incorporating prior knowledge and multi-kernel into linear programming support vector regression[J]. *Soft Computing*, 2015, 19(7): 2047-2061.
- [9] Shaobo L, Jinshan Z. Fast Learning With Polynomial Kernels[J]. *IEEE Transactions on Cybernetics*, 2018:1-13.
- [10] Boyd S, Vandenberghe L. *Convex optimization*[M]. Cambridge university press, 2004.
- [11] Maalouf M, Trafalis T B, Adrianto I. Kernel logistic regression using truncated Newton method[J]. *Computational management science*, 2011, 8(4): 415-428.
- [12] Platt J C. 12 fast training of support vector machines using sequential minimal optimization[J]. *Advances in kernel methods*, 1999: 185-208.
- [13] Keerthi S S, Duan K B, Shevade S K, et al. A fast dual algorithm for kernel logistic regression[J]. *Machine learning*, 2005, 61(1-3): 151-165.
- [14] Rossmann T. The average size of the kernel of a matrix and orbits of linear groups[J]. *Proceedings of the London Mathematical Society*, 2018, 117(3): 574-616.
- [15] Ubaru S, Saad Y, Seghouane A K. Fast estimation of approximate matrix ranks using spectral densities[J]. *Neural computation*, 2017, 29(5): 1317-1351.
- [16] Stražar M, Curk T. Learning the kernel matrix via predictive low-rank approximations[OL]. arXiv:1601.04366, 2016.
- [17] Phillips J M, Tai W M. Improved coresets for kernel density estimates[C]. *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*. Society for Industrial and Applied Mathematics, 2018: 2718-2727.
- [18] Needell D, Srebro N, Ward R. Stochastic gradient descent, weighted sampling, and the randomized Kaczmarz algorithm[J]. *Mathematical Programming*, 2016, 155(1-2):549-573.
- [19] Pourkamali-Anaraki F, Becker S. Randomized Clustered Nystrom for Large-Scale Kernel Machines[OL]. arXiv:1612.06470, 2016.
- [20] CHEN C, MATHEMATICS H H J S C. Global existence of real roots and random Newton flow algorithm for nonlinear system of equations [J]. 2017, 60(7): 1-12.
- [21] Xia J, Falco N, Benediktsson J A, et al. Hyperspectral image classification with

rotation random forest via KPCA[J]. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens, 2017, 10(4): 1601-1609.

[22] Zhu Z B, Song Z H. A novel fault diagnosis system using pattern classification on kernel FDA subspace[J]. Expert Systems with Applications, 2011, 38(6): 6895-6905.

[23] Ong C S, Smola A J, Williamson R C. Learning the kernel with hyperkernels[J]. Journal of Machine Learning Research, 2005, 6(Jul): 1043-1071.

[24] Rakotomamonjy A, Bach F R, Canu S, et al. SimpleMKL[J]. Journal of Machine Learning Research, 2008, 9(Nov): 2491-2521.

[25] Cortes C, Mohri M, Rostamizadeh A. Algorithms for learning kernels based on centered alignment[J]. Journal of Machine Learning Research, 2012, 13(Mar): 795-828.

[26] Lee J H, Lee G M. On optimality conditions and duality theorems for robust semi-infinite multiobjective optimization problems[J]. Annals of Operations Research, 2018, 269(1-2): 419-438.

[27] He L, Li Y, Zhang X, et al. Incremental spectral clustering via fastfood features and its application to stream image segmentation[J]. Symmetry, 2018, 10(7): 272-289.

[28] Song B, Li J, Dalla Mura M, et al. Remotely sensed image classification using sparse representations of morphological attribute profiles[J]. IEEE transactions on geoscience and remote sensing, 2014, 52(8): 5122-5136.

[29] Cortes C, Mohri M, Rostamizadeh A. Algorithms for learning kernels based on centered alignment[J]. Journal of Machine Learning Research, 2012, 13(Mar): 795-828.

[30] Fang L, Li S, Duan W, et al. Classification of hyperspectral images by exploiting spectral-spatial information of superpixel via multiple kernels[J]. IEEE transactions on geoscience and remote sensing, 2015, 53(12): 6663-6674.

[31] Li J, Marpu P R, Plaza A, et al. Generalized composite kernel framework for hyperspectral image classification[J]. IEEE transactions on geoscience and remote sensing, 2013, 51(9): 4816-4829.