

基于改进的Askey-Wilson 多项式核函数

张菁华, 闫焱*

华北理工大学理学院, 河北 唐山

收稿日期: 2022年5月10日; 录用日期: 2022年5月27日; 发布日期: 2022年6月13日

摘要

基于Askey-Wilson多项式的性质, 给出了两种Askey-Wilson多项式形式, 分别与Rbf核函数和三角核函数进行组合, 构造混合支持向量机AW核函数。所构造的AW核函数都含有一个重要核参数 q , 可通过改变 q 值来调整核函数。实验选取双螺旋线及6个UCI数据集, 比较AW核函数在数据集上的泛化性能和鲁棒性。

关键词

Askey-Wilson多项式, 三角核函数, Rbf核函数

Based on Improved Askey-Wilson Polynomial Kernel Function

Puhua Zhang, Yan Yan*

College of Science, North China University of Science and Technology, Tangshan Hebei

Received: May 10th, 2022; accepted: May 27th, 2022; published: Jun. 13th, 2022

Abstract

Based on the properties of Askey-Wilson polynomials, two forms of Askey-Wilson polynomials are given, which are combined with Rbf kernel function and triangular kernel function respectively to construct a hybrid support vector machine AW kernel function. The constructed AW kernel functions all contain an important kernel parameter q , which can be adjusted by changing the q value. The experiment selects the double helix and 6 UCI datasets to compare the generalization performance and robustness of the AW kernel function on the datasets.

*通讯作者。

Keywords

Askey-Wilson Polynomial, Trigonometric Kernel Function, Rbf Kernel Function

Copyright © 2022 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

一元 Askey-Wilson 多项式 $p_n(x; a, b, c, d | q)$ 是 n 阶的 x 多项式, 包含五个参数 a 、 b 、 c 、 d 和 q 。参考标准的基本超几何表示法[1]和该多项式的三项递推表达式关系, 可以推断 Askey-Wilson 多项式为正交多项式。正交多项式形成的正交函数系是分析数学中的一个重要组成部分, 正交多项式在实际应用中有着良好的鲁棒性, 现已被广泛应用于机器学习及数理统计等各方面。从分析组合角度出发, Askey-Wilson 多项式可以给出一个明确的绝对连续测度的 n 阶矩函数 $u_n(a, b, c, d; q)$, 该函数是包含参数 a 、 b 、 c 、 d 和 q 并混合二项式和 q -二项式的有理函数。Jang Soo Kim [2]通过组合模型的 n 阶矩函数, 给出了新的 n 阶 Askey-Wilson 多项式组合形式, 所包含的参数在多项式中权重有所不同, 当参数 $d = 0$ 时, 得到的 Askey-Wilson 多项式表达式为对称的多项式, 同样特殊情况 $b = -a$, $d = -c$ 也是如此。Askey-Wilson 多项式在正交多项式领域中是最新的研究成果, 构造支持向量机的核函数也是十分有效的。Luís Daniel Abreu [3]等人基于 Askey-Wilson 多项式构造出多项式再生核函数, Michael R. Hoare [4]等人研究了其转移矩阵。Askey-Wilson 多项式选取不同参数的值, 可以获取不同的 Askey-Wilson 多项式核函数形式, 但由于参数比较多, 所以必须要有一种高效的寻参机制和核选择机制, 利用自适应优化策略来控制参数。正交多项式核函数以往采用的是 Rbf 核函数与泛化正交多项式乘积的组合核。三角核函数有着与 Rbf 核函数相当的性能, 本文提出了 Askey-Wilson 多项式两类形式分别与 Rbf 核函数和三角核函数组合形成混合核函数, 并考察每一种混合核函数在支持向量机上的分类能力。

2. Askey-Wilson 核函数

2.1. Askey-Wilson 多项式

Askey-Wilson 多项式表达式[5]为:

$$p_n(x; a, b, c, d | q) = (ab, ac, ad; q)_n a^{-n} {}_4\Phi_3 \left(\begin{matrix} q^{-n}, abcdq^{n-1}, ae^{i\theta}, ae^{-i\theta} \\ ab, ac, ad \end{matrix} \middle| q; q \right),$$

其中 $x = \cos \theta$, 且 $|q| < 1$ 。

Askey-Wilson 多项式的递推公式[2]:

$$A_n p_{n+1}^\alpha(x) = (2x - B_n) p_n^\alpha(x) - C_n p_{n-1}^\alpha(x),$$

其中

$$A_n = \frac{a^{-1} (1 - abq^{n+\alpha}) (1 - acq^{n+\alpha}) (1 - adq^{n+\alpha}) (1 - abcdq^{n+\alpha-1})}{(1 - abcdq^{2n+2\alpha-1}) (1 - abcdq^{2n+2\alpha})}$$

$$C_n = \frac{a(1-bcq^{n+\alpha})(1-bdq^{n+\alpha})(1-cdq^{n+\alpha})(1-q^{n+\alpha})}{(1-abcdq^{2n+2\alpha-2})(1-abcdq^{2n+2\alpha-1})}$$

$$B_n = a + a^{-1} - A_n - C_n$$

Askey-Wilson多项式的正交关系[6]为:

$$\frac{1}{2\pi} \int_{-1}^1 \frac{\omega(x)}{\sqrt{1-x^2}} p_m(x; a, b, c, d | q) p_n(x; a, b, c, d | q) dx = h_n \delta_{mn}$$

其中 $h(x) = (e^{i\theta}, e^{-i\theta}; q)_\infty$, $\max(|a|, |b|, |c|, |d|, q) < 1$ 。

2.2. 三角核函数

三角核函数是根据 Laplace 核函数 $K(x, z) = \exp\left(-\frac{\|x-z\|}{\sigma}\right)$ 改进而来。利用泰勒展开式, 当 $\|x-z\| \rightarrow 0$ 时, $K(x, z)$ 近似等于 $1 - \frac{\|x-z\|}{\sigma}$ 。当 σ 为 1 或者 2 且 $z=0$, 固定 x 范围, $K(x, z) = 1 - \frac{\|x-z\|}{\sigma}$ 核函数关于 $x=0$ 是一个对称的三角图形, 故可称为三角核函数[7]。

2.3. 两类 AW 核函数

第一类 Askey-Wilson 多项式, $b=c=d=0$, $a=q$ 时, 根据递推公式 $A_n = q^{-1}$ 、 $C_n = q(1-q^n)$ 、 $B_n = q^{n+1}$, 1到3阶 AW 广义多项式为

$$p_1^1(x; q | q) = -q^2 + 2x$$

$$p_2^1(x; q | q) = (-q^2 + 2x) p_1^1(x; q | q)^T + q^2 - 1$$

$$p_3^1(x; q | q) = (-q^3 + 2x) p_2^1(x; q | q)^T + (q^3 - 1)(-q^2 + 2x)$$

x 为行向量, $p_n^1(x; q | q)^T$ 为 $p_n^1(x; q, q)$ 转置。

第二类 Askey-Wilson 多项式, $c=d=0$, $b=-a$, $a=q$ 时, 根据递推公式 $A_n = (1+q^{n+2})/q$ 、 $C_n = q(1-q^n)$ 、 $B_n = 0$ 1到3阶的 AW 广义多项式为

$$p_1^2(x; q | q) = 2qx / (q^2 + 1)$$

$$p_2^2(x; q | q) = 2qx p_1^2(x; q | q)^T / (1+q^3) + (q^3 - q^2) / (1+q^3)$$

$$p_3^2(x; q | q) = 2qx p_2^2(x; q | q)^T / (1+q^4) - (q^2 - q^4) p_1^2(x; q | q) / (1+q^4)$$

基于上述的两类 Askey-Wilson 多项式, 建立 n 阶 Askey-Wilson 高斯核函数和 Askey-Wilson 三角核函数

第一类 AW 高斯核函数 $K_{Gauss-H}^1(x, z) = \exp\left(-\frac{\|x-z\|^2}{d}\right) \sum_{j=0}^n P_n^1(x) (P_n^1(z))^T$, 第一类 AW 三角核函数

$K_{Tri-H}^1(x, z) = \left(1 - \frac{\|x-z\|}{\sigma_0}\right) \sum_{j=0}^n P_n^1(x) P_n^1(z)^T$ 。同理可以构建第二类 AW 高斯核函数 $K_{Gauss-H}^2(x, z)$ 和 第二类 AW 三角核函数 $K_{Tri-H}^2(x, z)$ 。

根据 Mercer 条件[6], 若 $\sum_{i=1}^n \left(\int p_i(x) \varphi(x) dx \right)^2 \geq 0$, 则 $K(x, z) = \sum_{i=1}^n p_i(x) (p_i(z))^T$ 是核函数。由于 $\exp\left(-\frac{\|x-z\|^2}{d}\right)$ 是 Rbf 核函数的一种形式, 所以核函数成立。三角函数 $1 - \frac{\|x-z\|}{\sigma_0}$, 基于其单调性, 利用 Mercer 定理, 证明也是一个核函数[6]。再根据核函数的性质, 核函数的乘积组合依旧是核函数, 所以第一类 AW 多项式混合函数和第二类 AW 多项式混合函数都可以作为支持向量机的核函数。

3. 实验结果

3.1. 在双螺旋分类对比

本文实验选择了 238 个样本点的双螺旋线数据集, 通过第一类 AW 高斯核函数、第一类 AW 三角核函数、第二类 AW 高斯核函数、第二类 AW 三角核函数以及 Rbf 核函数来进行对比。表 1 列出了这 5 种核函数在双螺旋集上训练精度为 100% 时的分类间隔及支持向量的个数。如下是各类 AW 核函数在训练集上的分类界面图。

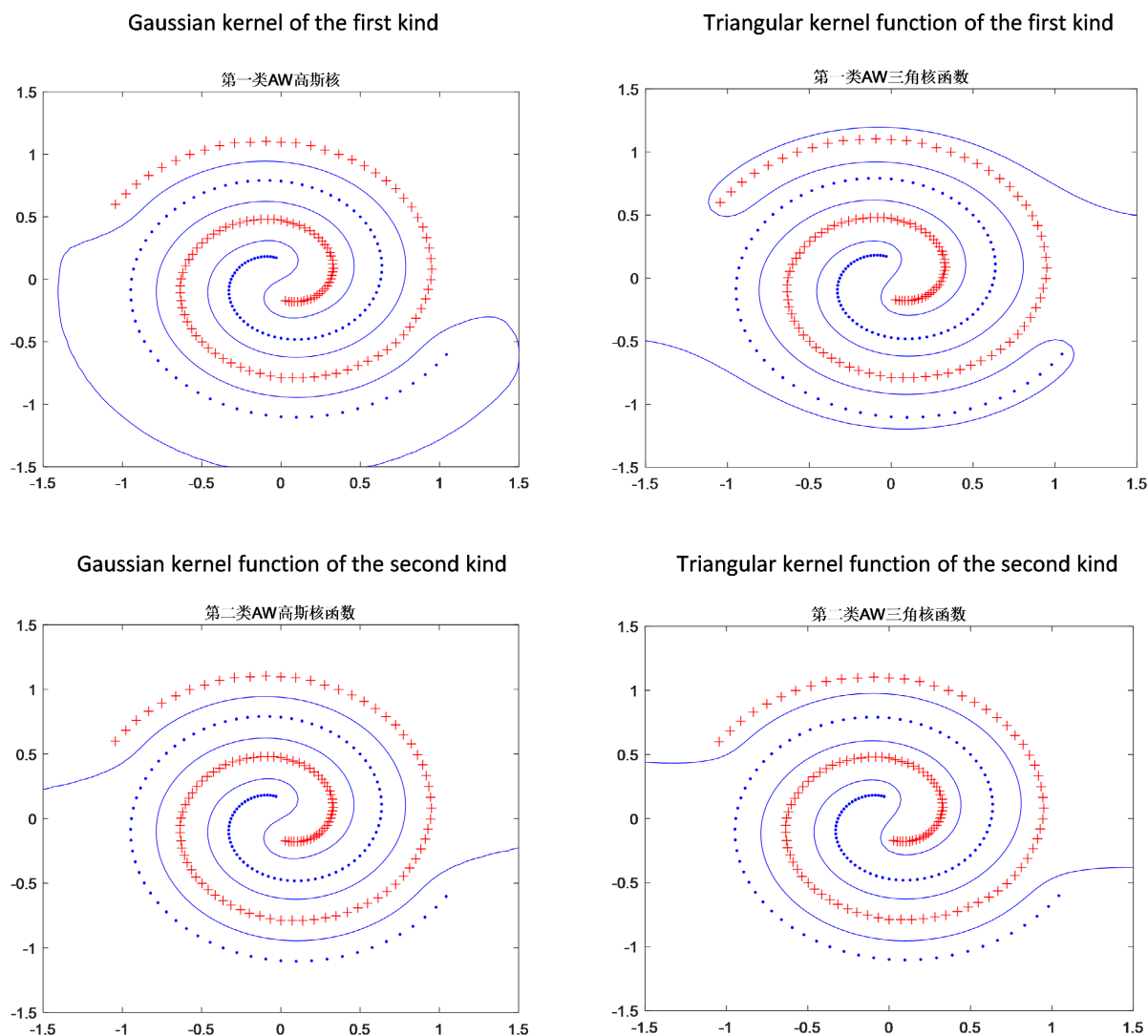


Table 1. Training results of kernel function on double helix**表 1.** 核函数在双螺旋线上训练结果

核函数	第一类 AW 高斯核	第一类 AW 三角核	第二类 AW 高斯核	第二类 AW 三角核	Rbf
参数	$n = 1$	$n = 1$	$n = 1$	$n = 1$	$\sigma = 0.1$
最大间隔	0.1608	0.16445	0.16999	0.17832	0.16006
SV 数	238	238	238	238	238

实验结果表明, AW 核函数都能将双螺旋数据集准确的分类, 且间隔都要优于 Rbf 核函数, 尤其是第二类 AW 三角核函数最大间隔最好, 其次是第二类 AW 高斯核函数。

3.2. 标准 UCI 数据集上的分类对比

选取的标准 UCI 数据集属性如表 2 所示。在这里只考虑 1 到 3 阶的 AW 核函数, 核参数 q 取不同值时, 1 到 3 阶 AW 核函数对数据预测会产生不同准确率。选取最高的准确率, 标明对应阶数, 则得到实验结果如表 3。根据表 3 的实验结果, 挑选出 AW 核函数在每个数据集上最好的分类效果, 然后和 Rbf 核函数、Poly 核函数分类效果做出对比, 得到表 4。

Table 2. Attributes of experimental UCI data sets**表 2.** 实验种 UCI 数据集的属性

数据集	训练集 样本数	测试集 样本数	维数	类别
vowel	176	352	10	11
sonar	70	138	60	2
segment	359	718	19	7
vechicle	282	564	18	4
svmguide2	131	260	20	3
liver_disorders	49	96	5	2

Table 3. Precision of four kinds of AW kernel function results: %**表 3.** 四类 AW 核函数结果精度: %

数据集	核函数类型	$q = -0.8$	$q = -0.6$	$q = -0.4$	$q = -0.2$	$q = 0.2$	$q = 0.4$	$q = 0.6$	$q = 0.8$	
vowel	第一类 AW 高斯核	80.1136 ($n = 3$)	80.3977 ($n = 3$)	80.9659 ($n = 3$)	78.4091 ($n = 1$)	80.3977 ($n = 2$)	82.1023 ($n = 1$)	85.5114 ($n = 2$)	82.3864 ($n = 3$)	
	第一类 AW 三角核	98.5795 ($n = 1$)	98.5795 ($n = 1$)	98.5795 ($n = 1$)	98.5795 ($n = 1$)	98.5795 ($n = 1$)	98.5795 ($n = 1$)	98.8636 ($n = 2$)	98.5795 ($n = 1$)	
	第二类 AW 高斯核	83.8068 ($n = 3$)	84.6591 ($n = 3$)	84.9432 ($n = 3$)	70.1705 ($n = 1$)	70.1705 ($n = 1$)	83.5227 ($n = 3$)	83.5227 ($n = 3$)	86.6477 ($n = 3$)	
	第二类 AW 三角核	95.7386 ($n = 1$)	95.1705 ($n = 1$)	95.7386 ($n = 2$)	89.7727 ($n = 2$)	89.7727 ($n = 2$)	95.7386 ($n = 2$)	97.7273 ($n = 3$)	98.0114 ($n = 2$)	
	sonar	第一类 AW 高斯核	83.3333 ($n = 3$)	81.8841 ($n = 1$)	82.6087 ($n = 3$)	86.9565 ($n = 3$)	87.6812 ($n = 2$)	87.6812 ($n = 3$)	85.5072 ($n = 3$)	86.2319 ($n = 2$)
		第一类 AW 三角核	83.3333 ($n = 1$)	84.058 ($n = 1$)	85.5072 ($n = 2$)	86.9565 ($n = 1$)	84.7826 ($n = 1$)	84.058 ($n = 1$)	84.7826 ($n = 2$)	84.7826 ($n = 2$)

Continued

	第二类 AW 高斯核	87.6812 (n = 3)	87.6812 (n = 3)	88.4058 (n = 1)	86.9565 (n = 1)	86.9565 (n = 1)	88.4058 (n = 1)	87.6812 (n = 3)	87.6812 (n = 3)
	第二类 AW 三角核	84.7826 (n = 3)	85.5072 (n = 3)	85.5072 (n = 3)	84.7826 (n = 2)	84.7826 (n = 2)	85.5072 (n = 2)	85.5072 (n = 3)	85.5072 (n = 3)
segment	第一类 AW 高斯核	92.0613 (n = 1)	89.8329 (n = 1)	90.8078 (n = 1)	91.6435 (n = 1)	91.7827 (n = 3)	91.6435 (n = 1)	91.6435 (n = 1)	92.2006 (n = 1)
	第一类 AW 三角核	95.4039 (n = 1)	95.2646 (n = 1)	95.4039 (n = 1)	95.5432 (n = 1)	95.6825 (n = 1)	95.6825 (n = 1)	95.6825 (n = 1)	95.6825 (n = 1)
	第二类 AW 高斯核	91.7827 (n = 1)	91.2256 (n = 1)	91.922 (n = 3)	92.0613 (n = 2)	92.0613 (n = 2)	92.0613 (n = 2)	91.3649 (n = 3)	91.7827 (n = 1)
	第二类 AW 三角核	94.7075 (n = 3)	94.7075 (n = 2)	94.7075 (n = 2)	93.1755 (n = 3)	93.1755 (n = 3)	94.9861 (n = 2)	94.9861 (n = 2)	94.8468 (n = 2)
vechicle	第一类 AW 高斯核	64.8936 (n = 1)	64.8936 (n = 3)	64.8936 (n = 1)	65.7801 (n = 3)	65.0709 (n = 2)	67.5532 (n = 3)	65.6028 (n = 3)	69.6809 (n = 2)
	第一类 AW 三角核	78.9007 (n = 1)	79.078 (n = 1)	79.078 (n = 1)	79.078 (n = 1)	79.2553 (n = 1)	79.4326 (n = 1)	79.4326 (n = 1)	79.6099 (n = 1)
	第二类 AW 高斯核	69.8582 (n = 3)	68.9716 (n = 3)	69.5035 (n = 3)	65.6028 (n = 3)	65.9574 (n = 3)	68.0851 (n = 3)	67.7305 (n = 3)	66.6667 (n = 3)
	第二类 AW 三角核	80.8511 (n = 2)	81.5603 (n = 2)	80.8511 (n = 3)	68.0851 (n = 3)	68.617 (n = 3)	80.1418 (n = 3)	81.2057 (n = 3)	80.4965 (n = 3)
svmguide2	第一类 AW 高斯核	75 (n = 1)	76.1538 (n = 1)	76.9231 (n = 1)	80 (n = 1)	80 (n = 1)	76.9231 (n = 1)	76.5385 (n = 1)	76.5385 (n = 3)
	第一类 AW 三角核	81.5385 (n = 1)	81.9231 (n = 3)	81.5385 (n = 1)	81.5385 (n = 1)	81.5385 (n = 1)	81.5385 (n = 1)	82.3077 (n = 2)	82.3077 (n = 3)
	第二类 AW 高斯核	61.5385 (n = 3)	56.9231 (n = 2)	56.5385 (n = 1)	56.9231 (n = 2)	56.9231 (n = 2)	56.5385 (n = 1)	56.9231 (n = 2)	61.5385 (n = 3)
	第二类 AW 三角核	81.9231 (n = 1)	81.9231 (n = 1)	81.5385 (n = 1)	81.5385 (n = 1)	81.5385 (n = 1)	81.5385 (n = 1)	81.9231 (n = 3)	81.9231 (n = 3)
liver_disorders	第一类 AW 高斯核	68.75 (n = 1)	68.75 (n = 1)	68.75 (n = 1)	67.7083 (n = 1)	66.6667 (n = 2)	70.8333 (n = 3)	68.75 (n = 2)	66.6667 (n = 3)
	第一类 AW 三角核	69.7917 (n = 3)	70.8333 (n = 2)	69.7917 (n = 2)	71.875 (n = 2)	71.875 (n = 2)	69.7917 (n = 2)	70.8333 (n = 2)	70.8333 (n = 2)
	第二类 AW 高斯核	68.75 (n = 1)	68.75 (n = 1)	68.75 (n = 1)	67.7083 (n = 2)	67.7083 (n = 1)	70.8333 (n = 2)	70.8333 (n = 2)	71.875 (n = 2)
	第二类 AW 三角核	70.8333 (n = 1)	70.8333 (n = 1)	70.8333 (n = 2)	67.7083 (n = 1)	67.7083 (n = 1)	71.875 (n = 3)	70.8333 (n = 2)	70.8333 (n = 1)

Table 4. Attributes of experimental UCI data sets**表 4.** 实验种 UCI 数据集的属性

	第一类 AW 高斯核	第一类 AW 三角核	第二类 AW 高斯核	第二类 AW 三角核	Rbf	Poly
vowel	85.5114	98.8636	86.6477	98.0114	59.0909	60.5114
参数	$n = 2, q = 0.6$	$n = 2, q = 0.6$	$n = 3, q = 0.8$	$n = 2, q = 0.8$	0.1	
sonar	87.6812	86.9565	88.4059	85.5072	84.7826	84.7826
参数	$n = 2, q = 0.2$	$n = 3, q = -0.2$	$n = 1, q = 0.4$	$n = 3, q = -0.6$	0.1	
segment	92.2006	95.6825	92.0613	94.9861	87.883	94.0111
参数	$n = 1, q = 0.8$	$n = 1, q = 0.2$	$n = 2, q = 0.2$	$n = 2, q = 0.4$	0.1	
vechicle	69.6809	79.6099	69.8582	80.8511	65.4255	69.3262
参数	$n = 2, q = 0.8$	$n = 1, q = 0.8$	$n = 3, q = -0.8$	$n = 2, q = -0.4$	0.1	
svmguid2	80	82.3077	61.5385	81.9231	56.5385	56.5385
参数	$n = 1, q = -0.2$	$n = 2, q = 0.6$	$n = 1, q = -0.8$	$n = 3, q = 0.6$	0.8	
liver_disorders	70.8333	71.875	71.875	71.875	57.75	66.6667
参数	$n = 3, q = 0.4$	$n = 2, q = -0.2$	$n = 2, q = 0.8$	$n = 3, q = 0.4$	0.1	

从表 4 可以看出, 第一类 AW 高斯核函数和第二类 AW 高斯核函数在 sonar 数据集上具有很好的表现, 在 vechicle 和 svmguid2 表现比第一类 AW 三角核函数, 第二类 AW 三角核函数要差一些。第一类 AW 三角核函数和第二类 AW 三角核函数可以通过调整核参数 q 的值从而获取最高的精度, 比高斯核函数和多项式核函数而言有明显优势, 因此有着非常好的鲁棒性的泛化能力。

参考文献

- [1] Gasper, G. and Rahman, M. (2004) Basic Hypergeometric Series. 2nd Edition, Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511526251>
- [2] Kim, J.S. and Stanton, D. (2014) Moments of Askey-Wilson Polynomials. *Journal of Combinatorial Theory, Series A*, **125**, 113-145.
- [3] Abreu, L.D. and Bouzeffour, F. (2010) A Paley-Wiener Theorem for the Askey-Wilson Function Transform. *Proceedings of the American Mathematical Society*, **138**, 2853-2862.
- [4] Hoare, M.R. and Rahman, M. (2008) A Probabilistic Origin for a New Class of Bivariate Polynomials. *SIGMA*, **4**, 89-107. <https://doi.org/10.3842/SIGMA.2008.089>
- [5] Ismail, M.E.H. and Stanton, D. (2015) Expansions in the Askey-Wilson Polynomials. *Journal of Mathematical Analysis and Applications*, **424**, 664-674. <https://doi.org/10.1016/j.jmaa.2014.11.048>
- [6] 田萌, 王文剑. 一类改进的埃尔米特核函数[J]. 计算机科学, 2014, 41(5): 239-242+274.
- [7] Fasshauer, G.E. and Hickernell, F.J. (2015) Solving Support Vector Machines in Reproducing Kernel Banach Spaces with Positive Definite Functions. *Applied and Computational Harmonic Analysis*, **38**, 115-139. <https://doi.org/10.1016/j.acha.2014.03.007>