

# Elastic-Net Regression Algorithm Based on Multi-Scale Gaussian Kernel

Yongli Xu<sup>#</sup>, Zhenjun Yang

Department of Mathematics, Beijing University of Chemical Technology, Beijing 100029, China

<sup>#</sup>Email: xuyongli2312@sina.com; buct\_yzj@126.com

## Abstract

This paper proposes an elastic-net algorithm based on multi-scale Gaussian kernels to deal with the approximation of regression function. We use Gaussian kernels with different kernel width to approximate the high and low frequency components of the regression function; then weighted L1 norm and L2 norm of the prediction function utilized as regularization term. In the simulation experiments, multiple Gaussian kernels based elastic-net obtains less prediction error and better sparse performance than single Gaussian kernel based elastic-net. In addition, multiple Gaussian kernels based elastic-net can precisely predict the high and low frequency components of objective function.

**Keywords:** Multi-kernel Based Elastic-net; Gaussian Kernel; Least Squares Regularized Regression; Sparsity

## 基于多尺度高斯核的弹性网回归算法\*

徐永利, 杨镇郡

北京化工大学 数学系, 北京 100029

**摘要:** 本文提出了一种基于多尺度高斯核的弹性网学习算法用以对回归函数的逼近。我们利用具有不同核宽度的高斯核函数构造基函数, 同时逼近目标函数的高频和低频成分。我们借鉴一般弹性网的思想, 取预测函数系数的 L1 范数和 L2 范数的加权组合作为最优化问题的正则项。在仿真数据和真实数据的实验中, 基于双高斯核的弹性网比单高斯核弹性网取得了更小的预测误差和更好的稀疏性表现。另外, 基于双高斯核的弹性网很好地预测了目标函数的高频和地频成分。

**关键词:** 多核弹性网; 高斯核; 最小二乘正则回归; 稀疏性

## 引言

Lasso 算法(least solute shrinkage and selection operator)和岭回归算法(ridge regression)是两种重要的正则化回归算法<sup>[2], [6]</sup>。Lasso 算法最早是作为高维统计的回归算法而提出来的, 其形式为:  $\|T-X\beta\|_2^2 + \lambda \sum_{j=1}^p |\beta_j|$ 。其最

优化的目标有两项: 残差平方和与回归系数绝对值之和。这里, 回归系数绝对值之和可以看做回归系数的 L1 范数。随着正则化参数  $\lambda$  的增大, 回归系数绝对值之和逐渐变小, 各回归系数相继收敛到 0。因此, Lasso 算法求得的系数具有稀疏性。如果 Lasso 算法中 L1 范数改为 L2 范数, 就变为岭回归算法。岭回归算法比 Lasso 算法具有更优的预测能力, 但求解系数的稀疏性不及 Lasso 算法。

Zou 和 Hastie 提出了一种兼具 Lasso 和岭回归算法优势的弹性网算法(elastic net)<sup>[2]</sup>, 其形式为:

$$\min_{(\beta_0, \beta) \in \mathbb{R}^{p+1}} R_\lambda(\beta_0, \beta) = \min_{(\beta_0, \beta) \in \mathbb{R}^{p+1}} \left[ \frac{1}{2N} \sum_{i=1}^N (y_i - \beta_0 - x_i^T \beta)^2 + \lambda P_\alpha(\beta) \right] \quad (1)$$

其中  $P_\alpha(\beta) = (1-\alpha) \frac{1}{2} \|\beta\|_2^2 + \alpha \|\beta\|_1$ 。弹性网的正则项是 Lasso 和岭回归正则项的凸组合。当  $\alpha=0$  时, 即为简单的岭回归; 当  $\alpha=1$  时, 即为 Lasso 算法。大量数值试验的结果表明, 弹性网算法同时具备良好的预测能力

\*基金资助: 受国家自然科学基金支持资助(11101024)。

和系数的稀疏性<sup>[3]</sup>。

核方法是机器学习的一种有效的方法<sup>[4]</sup>。核方法的主要思想是基于这样一个假设：在低维空间中不能线性分割的点集，通过转化为高维空间中的点集时，很有可能变为线性可分的支持向量机在理论和应用的发展，引起人们对核方法的高度关注<sup>[4]</sup>。核方法多是基于单个特征空间的单核方法。但是，当样本特征含有异构信息，样本规模很大，多维数据的不规则或数据在高维特征空间分布的不平坦，采用单个简单核进行映射的方式对所有样本进行处理并不合理<sup>[13]</sup>。

为了解决以上问题，大量关于核组合方法，即多核学习方法，相继被提出。这些方法期望通过多核的组合获得更优的预测性能。多尺度核方法是一种特殊的多核学习方法，将多个尺度的核进行融合。随着多尺度分析理论和小波理论的发展，多尺度核方法获得了较好的理论背景。多尺度核方法的基础就是要找到一组具有多尺度表示能力的核函数，在众多被广泛使用的核函数中，高斯核  $k(x,z)=\exp(-\frac{\|x-z\|^2}{2\sigma^2})$  是最常用的。因为它具有通用普遍的近似能力，并且也是一种典型的可多尺度化核。

本文提出一种基于多尺度高斯核的弹性网回归算法，并通过实验验证了这一算法在减小预测误差和提高预测函数系数的稀疏性方面的良好性能。

## 1 基于核的弹性网算法

回归问题是机器学习的一个重要课题。假设  $(X, d)$  是一个紧测度空间， $Y=\mathbb{R}$ 。假设  $\rho$  是  $Z=X*Y$  上的固定但未知的概率测度。回归函数定义为： $f_\rho(x)=\int_Y y d\rho(y|x)$ ， $x \in X$ ，这里  $\rho(y|x)$  是  $\rho$  在  $x$  诱导的条件概率。仅仅知道依据  $z=\{(x_i, y_i)\}_{i=1}^m$  选取的一组独立同分布的样本  $\rho$ ，我们的目标是学习一个函数  $f_z$ ，用已逼近  $f_\rho$ 。但是，利用系数的样本逼近一个函数是不是定的。对于此问题，正则化理论是一个有效的工具<sup>[11], [33]</sup>。

我们设计基于单高斯核的弹性网回归学习算法，定义为  $f(x)=\sum_{i=1}^m \beta_i K(x_i, x)$ ，这里  $(\beta_1, \dots, \beta_m)$  是下面最优化问题的解：

$$\min_{(\beta_1, \dots, \beta_m)} \left\{ \frac{1}{m} \sum_{j=1}^m \left( \sum_{i=1}^m \beta_i K(x_i, x_j) - y_j \right)^2 + \lambda P_\alpha(\beta) \right\}, \quad (2)$$

其中， $P_\alpha(\beta) = (1-\alpha) \frac{1}{2} \|\beta\|_2^2 + \alpha \|\beta\|_1$ ， $K(x,z)=\exp(-\frac{\|x-z\|^2}{2\sigma^2})$ ， $\lambda$  是正则化参数， $\alpha$  是正则化权重。最优化目标由经验风险和基于 L1 和 L2 系数的正则项组成。正则化参数  $\lambda$  起到平衡经验风险和正则项的作用。正则化权重  $\alpha$  起到平衡正则项中 L1 范数和 L2 范数的作用。当正则化权重  $\alpha=0$ ，正则项中只剩下 L2 范数，这时的弹性网算法即为基于核的岭回归算法；当正则化权重  $\alpha=1$ ，正则项中只剩下 L1 范数，这时的弹性网算法即为基于核的 LASSO 算法。

由优化问题（4）获得的预测函数，实际上是高斯核函数在以各样本点的输入  $x_i$  处的构成基函数的加权平均，这些基函数具有相同核宽度。然而在实际问题中，目标函数常常变化剧烈的高频成分和变化平缓的低频成分。在这种情况下，最优化问题(4)所对应的预测函数往往不能同时逼近目标函数的高低频成分。为此，我们设计基于多尺度高斯核的弹性网算法。为了简单起见，我们仅给出具有两个尺度的高斯核弹性网，具有多个尺度的高斯核弹性网可以类似的获得。基于双高斯核的弹性网算法定义为：

$$f(x) = \sum_{i=1}^m \beta_i K_1(x_i, x) + \sum_{i=1}^m \beta_{m+i} K_2(x_i, x) \quad (3)$$

其中  $K_1(x,z)=\exp(-\frac{\|x-z\|^2}{2\sigma_1^2})$  和  $K_2(x,z)=\exp(-\frac{\|x-z\|^2}{2\sigma_2^2})$  是两个具有不同核宽度的高斯核函数， $(\beta_1, \dots, \beta_{2m})$  是如下最优化问题的解：

$$\min_{(\beta_1, \dots, \beta_{2m})} \left\{ \frac{1}{m} \sum_{j=1}^m \left( \sum_{i=1}^m \beta_i K_1(x_i, x_j) + \sum_{i=1}^m \beta_{m+i} K_2(x_i, x_j) \right)^2 + \lambda P_\alpha(\beta) \right\}, \quad (4)$$

这里  $P_\alpha(\beta) = (1-\alpha)\frac{1}{2}\|\beta\|_2^2 + \alpha\|\beta\|_1$ ， $\lambda$  是正则化参数， $\alpha$  是正则化权重。最优化目标由经验风险和两个尺度的高斯核基于 L1 和 L2 系数的正则项组成。当正则化权重  $\alpha=0$ ，正则项中只剩下 L2 范数，这时的弹性网算法即为基于多尺度高斯核的岭回归算法；当正则化权重  $\alpha=1$ ，正则项中只剩下 L1 范数，这时的弹性网算法即为基于多尺度高斯核的 LASSO 算法。

## 2 实验设计与结果分析

本文通过一个仿真数据集和一个真实数据集对双高斯核弹性网算法进行分析，并与单高斯核弹性网算法进行对比。实验结果表明，双高斯核弹性网算法具有更好的稀疏性和更小的预测误差，并能很好的学习目标函数的高频和低频部分。

### 2.1 实验设计

对于每个模型，我们做 20 次重复实验。每次实验，生成  $3N$  个样本，并将其随机的平均分成三组。第一组样本（训练集）是用来获得基于高斯核的弹性网的解  $f$ ；第二组样本（验证集）是用来选择最优参数；最后一组样本（测试集）用于最后的测试。在我们的模型中，训练集、验证集和测试集的分布是相同的。训练

误差为  $e_{train} = \sqrt{\sum_{i=1}^N (y_i - \hat{y}_i)^2 / N}$ ；验证误差为  $e_{validation} = \sqrt{\sum_{i=N+1}^{2N} (y_i - \hat{y}_i)^2 / N}$ ；测试误差为  $e_{test} = \sqrt{\sum_{i=2N+1}^{3N} (y_i - \hat{y}_i)^2 / N}$ ，其中  $\hat{y}$  是对输出  $y$  的估计值。平均训练（验证、测试）误差为训练（验证、测试）误差 20 次重复实验的平均值。

对于优化问题（1）和（2），必须对参数  $\sigma$ 、 $\sigma_1$ 、 $\sigma_2$  和  $\lambda$  进行选择。在优化问题（2）中， $\sigma_1$  和  $\sigma_2$  表示高斯核的核宽度，表示正则化系数。由于优化问题（1）是（2）的特殊情况，我们给出（2）的参数选择方法，（1）的参数选择方法是类似的。给定  $\sigma_1$ 、 $\sigma_2$  和  $\lambda$ ，我们基于第一组样本求解优化问题（2）来获得函数  $f$ ，并且计算出它基于第二组样本的验证误差。不同的  $\sigma_1$ 、 $\sigma_2$  和  $\lambda$  产生不同的函数  $f$ ，最优的  $\sigma_1$ 、 $\sigma_2$  和  $\lambda$  是使验证误差最小的参数。所有的参数被取定后，预测函数是基于第一组样本的优化问题（2）的解  $f$ 。预测函数被用于预测新的样本的输出，并且由最后一组样本计算出平均测试误差。所有的实验都是用 MATLAB R2007b 在英特尔 2.5GHz CPU 下运行。为了比较单核弹性网和双核弹性网预测性能和稀疏性，我们用一个仿真数据和真实数据进行分析。

### 2.2 仿真数据实验

下面我们利用一个仿真数据来验证基于单高斯核和双高斯核弹性网算法的性能。首先我们如下构造一个目标函数：

$$y=5\sin(2x)+\sin(30x), 0 \leq x \leq \pi.$$

样本集为  $\{(x_i, f_1^*(x_i) + t\varepsilon_i)\}_{i=1}^N$ ，其中  $x_i \in (0, \pi)$  是均匀分布并且是独立的。噪声  $\varepsilon_i \in [-0.5, 0.5]$  是均匀分布并且是独立的。噪声的系数  $t \in \{0.1, 0.5\}$ 。惩罚因子  $a \in (i \times 10^{-1})_{i=0}^{10}$ 。样本数取为  $\{10, 20, 30, 40, 50, 60\}$ 。对于双核弹性网系数的选择，采用网格法， $\lambda_1, \lambda_2 \in \{10^{-8} + i \times 10^{-6}\}_{i=1}^{100}$ ， $\sigma_1, \sigma_2 \in \{i \times 10^{-2}\}_{i=1}^{10}$ 。为了减小计算复杂度，我们取  $\lambda_1 = \lambda_2$ 。对于单核弹性网系数的选择，同样采用网格法， $\lambda_1, \lambda_2 \in \{10^{-8} + i \times 10^{-6}\}_{i=1}^{100}$ ， $\sigma \in \{i \times 10^{-2}\}_{i=1}^{14}$ 。表 1 和表 2 分别给出了噪声系数数  $t$  取 0.5，正则化权重  $\alpha$  取 0，0.5，1 的情形下，基于单、双高斯核的弹性网算法预测函数的误差和 L1 范数。表 3 和表 4 分别给出了噪声系数数  $t$  取 0.1，正则化权重  $\alpha$  取 0，0.5，1 的情形下，基于单、双高斯核的弹性网算法预测函数的误差和 L1 范数。

表 1 单、双高斯核弹性网算法的误差对比（噪声  $t$  取 0.5，正则化权重  $\alpha$  分别取 0，0.5，1。样本数  $N$  分别去 10, 20, 30, 40, 50, 60）

| 样本                  | 10   | 20   | 30   | 40   | 50   | 60   |
|---------------------|------|------|------|------|------|------|
| 单核弹性网, $\alpha=0$   | 2.36 | 2.06 | 1.34 | 1.14 | 0.75 | 0.50 |
| 双核弹性网, $\alpha=0$   | 2.40 | 1.79 | 1.09 | 0.90 | 0.61 | 0.43 |
| 单核弹性网, $\alpha=0.5$ | 2.31 | 2.06 | 1.31 | 1.13 | 0.73 | 0.51 |
| 双核弹性网, $\alpha=0.5$ | 2.41 | 1.76 | 1.11 | 0.92 | 0.61 | 0.43 |
| 单核弹性网, $\alpha=1$   | 2.31 | 2.05 | 1.29 | 1.12 | 0.74 | 0.51 |
| 双核弹性网, $\alpha=1$   | 2.42 | 1.75 | 1.13 | 0.91 | 0.59 | 0.42 |

表 2 单、双高斯核弹性网算法的 L1 范数对比 (噪声  $t$  取 0.5, 正则化权重  $a$  分别取 0, 0.5, 1。样本数  $N$  分别去 10,20,30,40,50,60)

| 样本                  | 10    | 20     | 30     | 40     | 50     | 60     |
|---------------------|-------|--------|--------|--------|--------|--------|
| 单核弹性网, $\alpha=0$   | 41.05 | 104.21 | 185.60 | 309.18 | 287.50 | 212.71 |
| 双核弹性网, $\alpha=0$   | 34.92 | 84.44  | 121.49 | 171.27 | 251.57 | 200.18 |
| 单核弹性网, $\alpha=0.5$ | 35.28 | 96.91  | 181.33 | 299.91 | 278.33 | 190.53 |
| 双核弹性网, $\alpha=0.5$ | 34.08 | 78.46  | 122.97 | 159.04 | 240.53 | 185.88 |
| 单核弹性网, $\alpha=1$   | 35.03 | 95.08  | 180.72 | 284.74 | 262.27 | 177.92 |
| 双核弹性网, $\alpha=1$   | 33.29 | 70.76  | 130.98 | 152.75 | 179.60 | 169.89 |

表 3 单、双高斯核弹性网算法的误差对比 (噪声  $t$  取 0.1, 正则化权重  $a$  分别取 0, 0.5, 1。样本数  $N$  分别去 10,20,30,40,50,60)

| 样本                  | 10   | 20   | 30   | 40   | 50   | 60   |
|---------------------|------|------|------|------|------|------|
| 单核弹性网, $\alpha=0$   | 2.43 | 1.94 | 1.53 | 0.89 | 0.69 | 0.51 |
| 双核弹性网, $\alpha=0$   | 2.38 | 1.56 | 1.26 | 0.79 | 0.56 | 0.42 |
| 单核弹性网, $\alpha=0.5$ | 2.50 | 1.95 | 1.49 | 0.88 | 0.70 | 0.51 |
| 双核弹性网, $\alpha=0.5$ | 2.38 | 1.55 | 1.27 | 0.80 | 0.56 | 0.42 |
| 单核弹性网, $\alpha=1$   | 2.46 | 1.92 | 1.47 | 0.90 | 0.70 | 0.51 |
| 双核弹性网, $\alpha=1$   | 2.40 | 1.58 | 1.27 | 0.81 | 0.56 | 0.43 |

表 4 单、双高斯核弹性网算法的 L1 范数对比 (噪声  $t$  取 0.1, 正则化权重  $a$  分别取 0, 0.5, 1。样本数  $N$  分别去 10,20,30,40,50,60)

| 样本                  | 10    | 20     | 30     | 40     | 50     | 60     |
|---------------------|-------|--------|--------|--------|--------|--------|
| 单核弹性网, $\alpha=0$   | 45.07 | 110.26 | 244.77 | 255.64 | 358.83 | 307.32 |
| 双核弹性网, $\alpha=0$   | 37.71 | 62.79  | 121.93 | 145.16 | 181.83 | 176.78 |
| 单核弹性网, $\alpha=0.5$ | 44.82 | 109.92 | 226.14 | 247.58 | 349.86 | 293.54 |
| 双核弹性网, $\alpha=0.5$ | 37.04 | 60.96  | 113.90 | 132.05 | 166.34 | 165.09 |
| 单核弹性网, $\alpha=1$   | 44.10 | 165.17 | 213.65 | 248.36 | 311.82 | 288.85 |
| 双核弹性网, $\alpha=1$   | 36.06 | 58.70  | 110.73 | 124.02 | 157.27 | 149.17 |

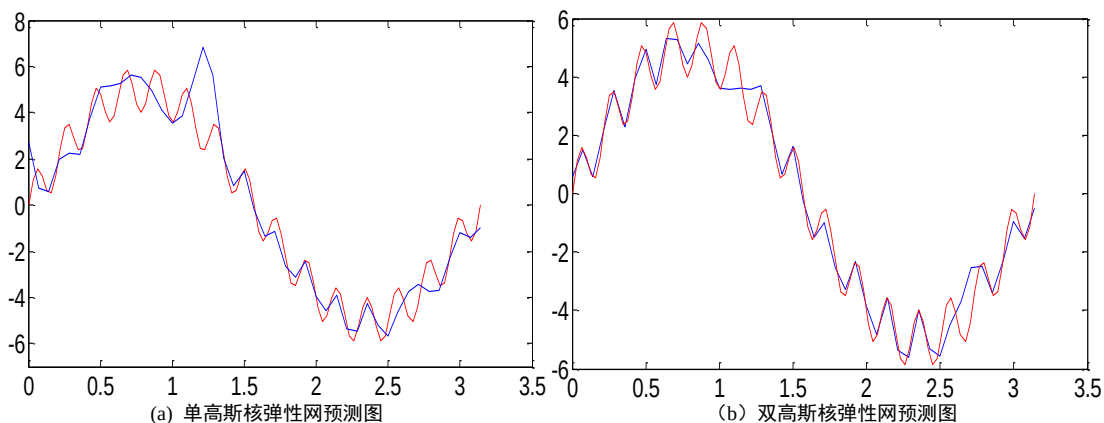


图 1 单、双高斯核弹性网算法的预测情况 (样本数  $N$  取 45, 红色曲线代表人工函数, 蓝色曲线代表单高斯核弹性网算法的预测函数函数, 绿色曲线为双高斯核弹性网算法的预测函数)

从表 1 和表 3 中可以看到, 在不同的噪声条件下, 随着样本数的增加, 单双核弹性网的误差都会逐渐减小。在相同的惩罚因子、噪声条件和样本数目条件下, 双高斯核弹性网比单高斯核弹性网具有更小的预测误差。另外, 随着正则化权重的增加, 单、双高斯核弹性网的平均预测误差没有明显的变化。从表 2 和表 4 可以看到, 在不同的噪声条件下, 随着样本数的增加, 单双核弹性网预测函数的 L1 范数都会先增大后减小。

在相同的正则化权重、噪声条件和样本数目条件下，双高斯核弹性网比单高斯核弹性网的预测函数具有更小的 L1 范数。另外，随着正则化权重的增加，单、双高斯核弹性网预测函数的 L1 范数都具有减小的趋势。图 1 给出了当样本数目为 45，正则化权重取 0.5，噪声系数取 0.5 时，单、双高斯核弹性网对目标函数的逼近情形。由于双核弹性网的预测函数由两个尺度的高斯核构成，我们可以把这个预测函数分为两个成分，分别用以预测目标函数的高频和低频部分。图 2 给出了当样本数取 200，在不同的正则化权重条件下，双核弹性网的预测函数的两个尺度成分对目标函数高低频成分的预测情况。从图 2 可以看到，当  $\alpha=0$ ，双核弹性网对目标函数的高低频成分的预测都有较大的偏差。当  $\alpha=1$ ，低频成分得到较好的预测，高频成分仍具有一些偏差。当  $\alpha=0.4$ ，目标函数的高低频成分都得到了很好地预测。

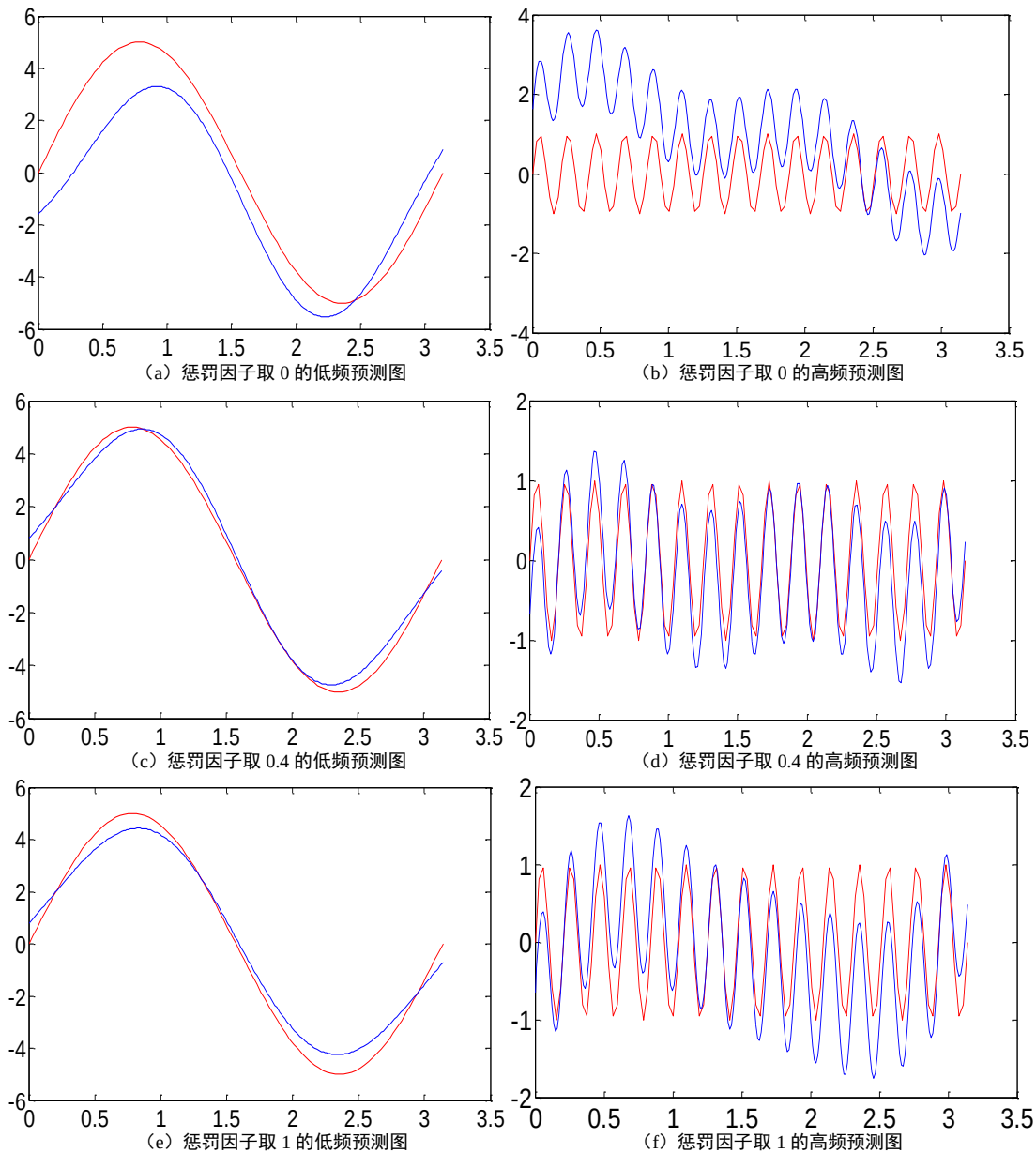


图 2 双高斯核弹性网算法对人工函数的高频分量和低频分量的预测（样本数  $N$  取 250，噪声系数  $t$  取 0.1。红色曲线代表人工函数分量，蓝色曲线代表预测函数分量）

## 2.3 真实数据实验

接下来，我们用一个真实的风速数据来验证基于高斯核的弹性网的预测性能。这个风速数据由 600 个时刻点的风速记录值组成（如下面图 3 所示）。

我们做 20 次实验，每次实验随机的将样本集平均分成三组，分别为训练集，验证集和测试集。对于系数的选择，用网格法。然后，分别用单双高斯核弹性网算法获得预测函数。对于参数选择我们采用网格化方法。其中，对于双高斯核弹性网算法，正则化参数和核宽度分别取  $\lambda_1, \lambda_2 \in \{10^{-10} + i \cdot 10^{-6}\}_{i=1}^{100}$ ， $\sigma_1, \sigma_2 \in \{2^i\}_{i=-1}^5$ 。对于单高斯核弹性网，正则化参数和核宽度分别取  $\lambda_1, \lambda_2 \in \{10^{-10} + i \cdot 10^{-6}\}_{i=1}^{100}$ ， $\sigma \in \{2^i\}_{i=-1}^5$ 。表 5 给出了在不同正则化权重条件下，单、双高斯核弹性网算法的平均预测误差和预测函数的平均 L1 范数。从表 5 可以看到，在相同的正则化权重条件下，双高斯核弹性网算法比单高斯核弹性网获得了更小的平均预测误差和更小的预测函数的 L1 范数。随着正则化权重的增加，单、双高斯核弹性网算法的平均预测误差没有明显的变化，预测函数的平均 L1 范数大致呈现下降的趋势。

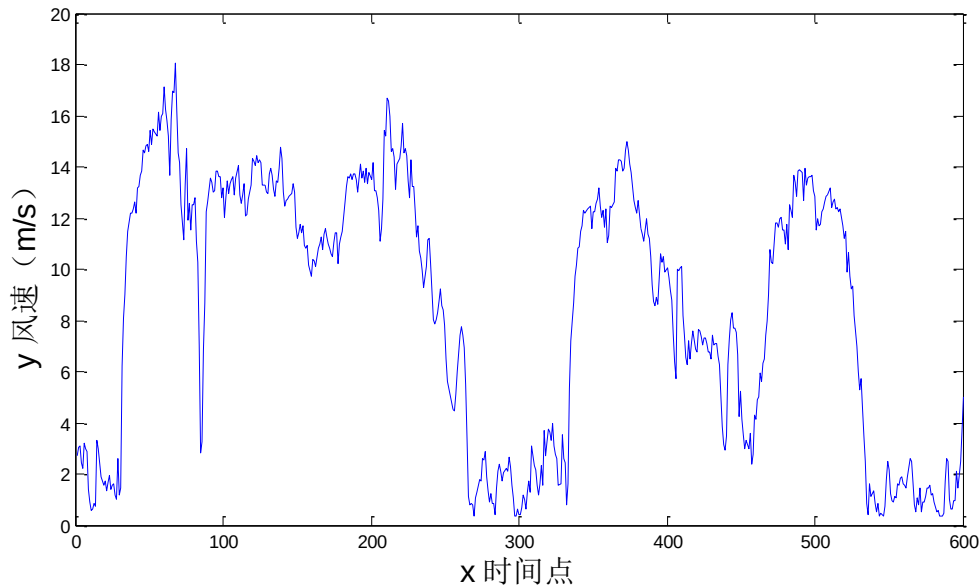


图 3 风速变化情况

表 5 单、双高斯核弹性算法网算法风速数据的误差和 L1 范数对比

| a   | 单核平均误差 | 双核平均误差 | 单核平均 L1 范数 | 双核平均 L1 范数 |
|-----|--------|--------|------------|------------|
| 0   | 1.386  | 1.289  | 3092       | 452        |
| 0.1 | 1.386  | 1.288  | 3051       | 449        |
| 0.2 | 1.386  | 1.287  | 3010       | 446        |
| 0.3 | 1.386  | 1.286  | 2960       | 443        |
| 0.4 | 1.386  | 1.286  | 2913       | 440        |
| 0.5 | 1.386  | 1.286  | 2873       | 437        |
| 0.6 | 1.387  | 1.285  | 2826       | 434        |
| 0.7 | 1.387  | 1.284  | 2776       | 432        |
| 0.8 | 1.387  | 1.284  | 2726       | 430        |
| 0.9 | 1.387  | 1.284  | 2679       | 427        |
| 1   | 1.388  | 1.283  | 2635       | 425        |

### 3 结论

本文提出高斯核双核弹性网算法，因为双核弹性网可以调节两个核宽度，所以能很好的拟合高频和低频的混合函数。通过与单核弹性网的对比，证明了上述理论并且很明显看到双核弹性网的优势。同时弹性网能吸收 loss 和宽回归算法的优点，平衡两种算法的缺点，即能表现稀疏性又能很好的学习函数的细节部分（高频部分和低频部分）。本文的算法改进了原有的算法，具有一定的意义。

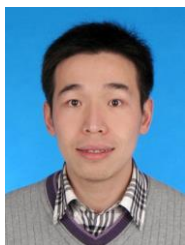
本文用了人工函数的数据和现实生活数据来验证理论具有一定的说服意义，数据都是真实有效的。

### REFERENCES

- [1] Yong-Li Xu, Di-Rong Chen, Han-Xiong Li. Least Square Regularized in Sum Space[J], IEEE TRANSACTION NETWORKS AND

- [2] De Mol, C.De Vito, E.Rosasco, L,et al. Elastic-Net Regularization in Learning Theory[J].2008,.DOI:
- [3] Kui Xiang,Bing-nan Li. Sparsity in Principal Component Analysis:A Survey[J]. Acta Electronica Sinica, 2012, 40(12):2525-2532. DOI:10.3969/j.issn.0372-2112.2012.12.027
- [4] Hong-Qiao Wang, Yan-Ning Cai, Fu-Chun Sun, Zong-Tao Zhao. Adaptive Sequence Learning and Applications for Multi-Scale Kernel Method [J]. PATTERN RECOGNITION AND ARTIFICIAL INTELLIGENCE, YR 2011, VO 24, IS 1: OP 72-81
- [5] Yan-Jun, Jue Wang. A Bi-Sparse Relational Learning Algorithm Based on Multiple Kernel Learning[J]. JOURNAL OF COMPUTER RESEARCH AND DEVELOPMENT,YR 2010, VO 47, IS 8, OP1400
- [6] 张勇进.带弹性网的稀疏主成分分析[D].华中科技大学,2006.DOI:10.7666/d.d047298
- [7] 付光辉.高维的强相关数据的模型选择[D].中南大学,2011.DOI:10.7666/d.y1918551
- [8] Chun-Ying Liu, Yan-Xiong Fang, Yu-Mian Yu. Research into Simultaneous Determination of Five Components by Ridge Regression Spectrophotometry[J].SPECTROSCOPY AND SPECTRAL ANALYSIS, 1999,19(4):629-631
- [9] Jerome Friedman,Trevor Hastie,Rob Tibshirani,et al. Regularization Paths for Generalized Linear Models via Coordinate Descent [J]. Journal of Statistical Software,2010,33(01)

### 【作者简介】



<sup>1</sup>徐永利（1982-），男，汉族，博士，讲师，机器学习、系统识别和生物医学图像分析，北京航空航天大学学士学位和博士学位。Email: xuyongli2312@sina.com



<sup>2</sup>杨镇郡（1991-），男，汉族，本科，学生，现就读北京化工大学。  
Email: buct\_yzj@126.com