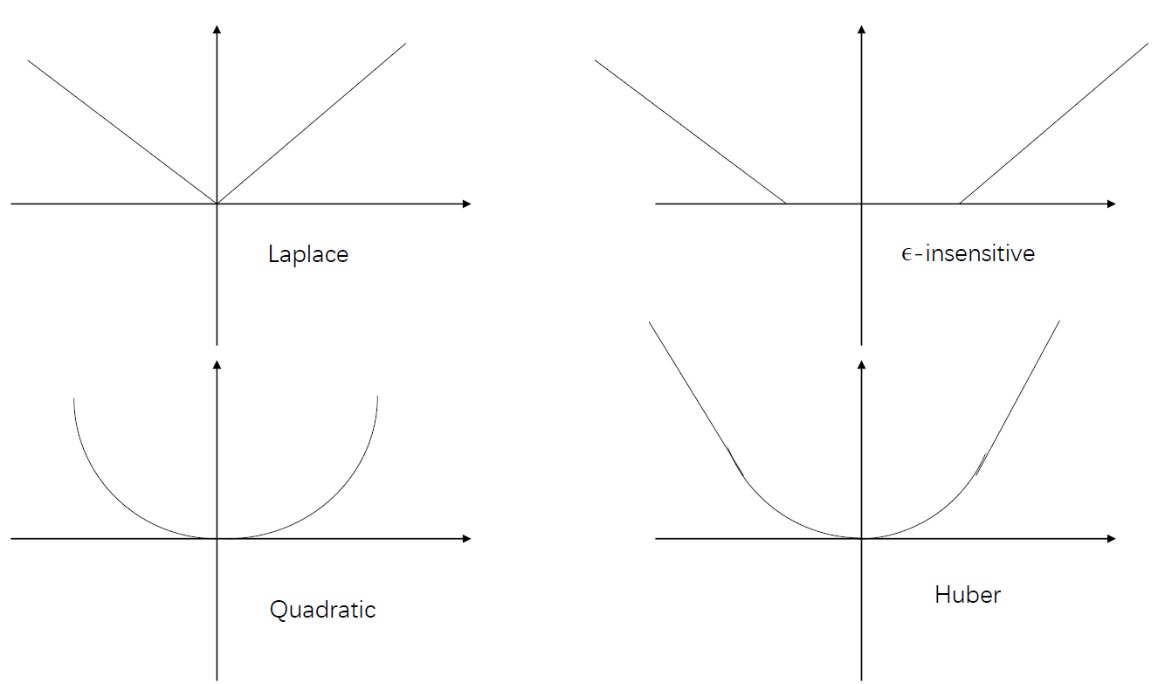
1. 支持向量回归Support Vector Regression回顾

考虑输入样本，其中并且作为相应的输出。令作为一种函数映射，将维向量空间的变量映射到希尔伯特空间，支持向量回归问题求解下面的公式：

1.1

其中，表示损失函数，而是正则系数。损失函数是可容忍拟合误差的一种度量，即对进行变换，常用的有Laplace损失，-insensitive损失，Quadratic损失，Huber损失等:

这一类损失的特点是对称性损失，我们选择Laplace损失，并将得到下面的形式：

1.2

这个式子的拉格朗日形式为：

其中，是非负拉格朗日乘子，并推导出它的对偶形式为：

1.4

其中，是一个对称正定矩阵，其中每个位置对应的元素为。最优解对应原始变量primal variables的形式为：

1.5

这样得到预测函数：

1.6

综合上述情况，支持向量回归SVR不需要明确定义映射函数，只需要核函数作为内积函数即可。引入如下的核函数：

1. Polynomials of degree d：
2. Polynomials of degree up to d:
3. Radial Function：
4. Polynomial Function：
5. Sigmoid Function:

其中是核函数的用户自定义可变参数。

1. 多核函数Multiple Kernel Function扩展形式

支持向量回归需要选择合适的核函数，如果有一组，即有一组特征映射函数**,**，如果利用所有的映射特征，等同于使用累加核函数，则目标函数转变为：

2.1

并有相应的对偶形式：

2.2

求解这个标准形式的优化方法有QP方法和SMO方法等。上式的求解，最终得到对偶参数的最优值，类似于（1.5）的参数表示为:

2.3

引入L1模正则化系数做尺度变换且。并定义每个参数，进而我们有如下的目标函数：

;

2.4

满足2.4式的最优参数 以及给出预测函数:

2.5

其中。求解2.2获得估计，再求解2.4获得以及，被称为多核学习的两步法Two-step method。两步法的第一步是样本点选择，第二步是核函数选择。值得注意的是，多核组合仍然满足mercer定理，可用于支持向量回归问题的内积核。

1. 数值分析与迭代算法

上面目标函数2.4的拉格朗日形式为：

3.1

上式有拉格朗日乘子，令

3.2

则有Karush-Kuhn-Tucker（KKT）条件为：

3.3

3.4

3.5

3.6

由可将松弛变量条件写为：

3.7

3.8

3.9

3.10

3.11

3.12

上面的3.5 至3.12且求解最小最大化的拉格朗日目标函数：

如果，由3.8则最大化目标函数必然有，同样由3.7则最小化目标函数必然有，否则目标函数都会趋于无穷大。根据3.6可得，进而由3.11得到，根据3.10可得，进而由3.5得到。因而则有这样的分析方法给出如下关系：

3.13

基于这三个条件，我们构建四个事件集合：

右侧划分Right

肘部位置Elbow

左侧划分Left

动态核集合 3.14

下面的原始系统Primal System成立：

3.15

3.16

如果样本点或者则有确定，由3.3可以得到：

3.17

因为如果满足，则由3.12可得，进而由3.4有：

， 3.18

我们称3.17和3.18为对偶系统。对于任何给定下满足KKT条件的解集合，必然满足3.15—3.18，也即原始系统和对偶系统成立，反之亦然。对于线性方程系统，原始系统有变量需要确定，也即,来满足个方程，而对偶系统有个变量需要确定，也即来满足个方程。为了同时满足两个系统,需要满足，因而，我们考虑如何进行集合初始化和遍历集合路径状态。

* 时的集合初始化，

初始时有唯一解，根据2.4则必然有。进而得到集合的初始状态，此时对输入序列进行降序排列得到，则有：

3.19

其中是向上取整函数,而表示在原始序列中的下标。根据3.14关于的定义，可以得到相应的集合划分，有集合的基数为,,。若，则有；若，则有；，，由3.3如果为偶数则否则；

初始化集合还需要向添加一个元素,我们考虑，事实上由系统的连续性，对于在0右邻域内的取值将不会改变集合,因而不发生变化，添加任意的，3.15-3.18都成立，因而我们只需要考虑满足2.4的最优以及参数的值。由优化目标函数：

3.20

上式等价于：

3.21

即有：

3.22

则有并且。

* 变动的临界值，变为

由线性系统的连续性在,确定下有，我们分析处的连续变化，即有：

3.23

因为，和是常数，并可以被系统线性系统唯一确定（非奇异）。因而，考虑处的线性增量，可以得到参数关于的线性函数：

,

3.24

可以确定对于较小的增量，3.24仍然满足,的集合划分条件不变，因而可以根据3.24式来更新系统参数。但存在一个临界值，使得原先的集合划分在3.24式下发生变化，我们将这种变化分为两个事件：

Event1：一个点离开或者，并且加入；即有一个点的残差从非0变为0；Event2：一个激活系数变为非激活，即从改变为；

3.24的线性方程保证了已经放入的点不会从集合里出来，但随着的线性增长，一定会恰好发生Event1或者Event2。

因而我们枚举出所有的事件临界点，对于所有的点，计算Event1的点：

对于所有的，计算Event2的点：

3.26

线性增长发生的事件，最早将在如下取值出现：

3.27

基于，我们可以更新得到变量。在给定第步参数和集合状态情况下，线性系统的关系确立，增长，发生Event1则增加1，发生Event2则减1，并将新的集合状态作为步参数。

* 变动的临界值，由变为

考虑对偶系统中作为的线性函数：

3.28

；其它的个变量可以被上述线性系统个等式唯一确定，因为此时满足。为了满足条件，我们需要进行“集合的点删除和集合的点加入点”操作，来改变集合状态。并分为两个事件：

Event3：一个点移除出集合，即将放入到或者集合中；

Event4：一个点加入到中；

我们让,因而3.17和3.18在集合状态不发生变化下恒成立。其中，Event3发生的临界点（）：

Event4发生需要满足3.18式成立,则有：

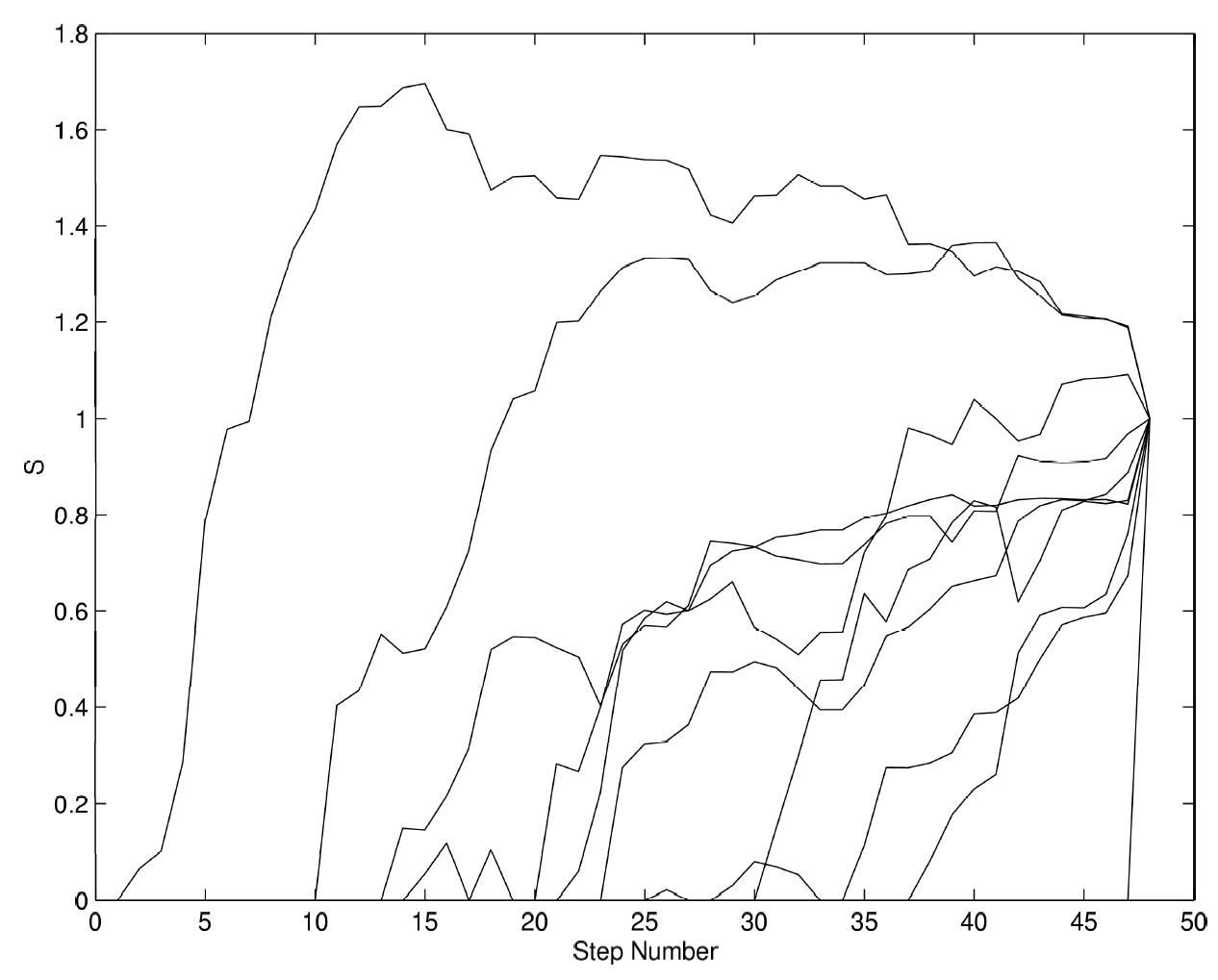
对偶系统提供了变动的最大临界值：

3.31

进而更新和对偶参数。并得到,，如果发生Event3则减1，发生Event4则加1。

* 迭代算法分析

上述两个步骤，其中一步是让线性增长，另一步使得线性递减，算法循环两个步骤直到减少到0停止。在两个相继发生的事件Event1/Event2到Event3/Event4之间，最优解与有着固定的线性函数关系。因而，最优解是的逐段线性函数，而每个路径拐角处则对应一个新事件。



图展示了50个样本10个备选径向基函数核的选择过程

**Initialization:** Calculate and ; determine ,and set ;

**Step1**：Solve the primal system 3.23；

**Step2**：Calculate Δ；identify the next event for the primal system and update the primal variables；

**Step3**：Solve the dual system 3.28

**Step4**：Calculate Δ; identify the next event for the dual system and update the dual variables;

**Step5**: If the stopping criterion is met, then stop the algorithm;

**Step6**: Otherwise, let and goto Step1.