* 1. First, note that multiplication of a Kernel by a positive constant yields a new kernel:

Note that each of the new bracketed terms defines a new mapping function equaling a new Kernel . The summing of two kernels should also yield a new Kernel and indeed:

* 1. . Additionally, , being a kernel, is positive semi definite -

.

then:

For to be a kernel, we must require it be positive semidefinite:

But since , this implies

That is, is the Zero kernel. But was explicitly said to not be the Zero kernel. By contradiction, is not a kernel.

* 1. Let .  
     Due to the conclusion from the recitation, the probability for a randomly drawn Dichotomy on a dimension N, with number of data points K, assuming (otherwise the probability is 1), is:

,  
which, in our case, translates to

.

Requiring , we get

Calculating numerically, we get:

The rest of the terms in the range can be clumped according to their thresholds

p requiredDim

0.001 77

0.002 78

0.003 79

0.005 80

0.007 81

0.011 82

0.016 83

0.023 84

0.032 85

0.044 86

0.059 87

0.078 88

0.1 89

0.128 90

0.16 91

0.197 92

0.239 93

0.285 94

0.335 95

0.388 96

0.443 97

0.499 98

0.556 99

0.611 100

0.664 101

0.714 102

0.76 103

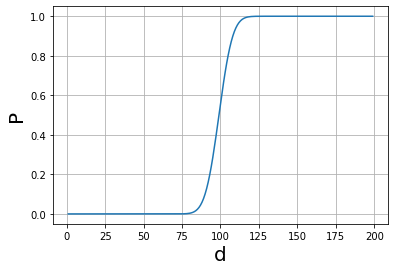
0.802 104

0.839 105

0.871 106

0.899 107

0.9 108

* 1. 

1. We require a mapping

Consider s.t. . As a result,

Observe that

And therefore:

Hence

As required. Thus is a kernel.

then the Lagrangian is  
 *,*and its gradient is given by

Requiring

From we have , thus, denoting and , we have:

And therefore:

Similarly, the other two permutations yield local extrema at

Substituting all 6 possible extrema back in :

So are maxima and are minima

Draw random points to form the dataset D. WLOG Label the points 1 if inside the cylinder and 0 if outside. Define

Define

Define

Define

Return ()

Denote a cylinder (necessarily contained within the concept) such that the probability (not necessarily volume!) difference between it and the concept equals some ε. It has a radius , and boundaries , . Divide the volume of the cylinder into three regions s.t

The probability that none of the m data points drawn, fall within these regions is bound by:

Requiring our confidence level to satisfy, m must satisfy

The sample complexity is thus

Time complexity: as all steps can be computed in linear time

|  |
| --- |
| Step |
| Draw m points |
| Label and group m points |
| Find |
| Find |
| Find |