RBF Space

Mahdi Akbari Zarkesh
9612762638

The Radial Basis Function Kernel

The **Radial basis function kernel**, also called the **RBF kernel**, or **Gaussian kernel**, is a kernel that is in the form of a radial basis function (more specifically, a Gaussian function). The RBF kernel is defined as

$$K_{\text{RBF}}(\mathbf{x}, \mathbf{x}') = \exp\left[-\gamma \left\|\mathbf{x} - \mathbf{x}'\right\|^2\right]$$

where γ is a parameter that sets the "spread" of the kernel.

The RBF kernel as a projection into infinite dimensions

Recall a kernel is any function of the form:

$$K(\mathbf{x}, \mathbf{x}') = \langle \psi(\mathbf{x}), \psi(\mathbf{x}') \rangle$$

where ψ is a function that projections vectors $\mathbf{x}$ into a new vector space. The kernel function computes the inner-product between two projected vectors.

As we prove below, the ψ function for an RBF kernel projects vectors into an infinite dimensional space. For Euclidean vectors, this space is an infinite dimensional Euclidean space.

That is, we prove that

$$\psi_{\text{RBF}} : \mathbb{R}^n \to \mathbb{R}^\infty$$

Proof:

Without loss of generality, let $\gamma = \frac{1}{2}$.

$$
\begin{aligned}
K_{\mathrm{RBF}}(\mathbf{x}, \mathbf{x}') &= \exp\left[-\frac{1}{2}\,||\mathbf{x} - \mathbf{x}'||^2\right] \\
&= \exp\left[-\frac{1}{2}\langle \mathbf{x} - \mathbf{x}', \mathbf{x} - \mathbf{x}' \rangle\right] \\
&= \exp\left[-\frac{1}{2}(\langle \mathbf{x}, \mathbf{x} - \mathbf{x}' \rangle - \langle \mathbf{x}', \mathbf{x} - \mathbf{x}' \rangle)\right] \\
&= \exp\left[-\frac{1}{2}(\langle \mathbf{x}, \mathbf{x} - \mathbf{x}' \rangle - \langle \mathbf{x}', \mathbf{x} - \mathbf{x}' \rangle)\right] \\
&= \exp\left[-\frac{1}{2}(\langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{x}' \rangle - \langle \mathbf{x}', \mathbf{x} \rangle + \langle \mathbf{x}', \mathbf{x}' \rangle)\right] \\
&= \exp\left[-\frac{1}{2}(||\mathbf{x}||^2 + ||\mathbf{x}'||^2 - 2\langle \mathbf{x}, \mathbf{x}' \rangle)\right] \\
&= \exp\left[-\frac{1}{2}\,||\mathbf{x}||^2 - \frac{1}{2}\,||\mathbf{x}'||^2\right] \exp\left[-\frac{1}{2} - 2\langle \mathbf{x}, \mathbf{x}' \rangle\right] \\
&= C e^{\langle \mathbf{x}, \mathbf{x}' \rangle} && C := \exp\left[-\frac{1}{2}\,||\mathbf{x}||^2 - \frac{1}{2}\,||\mathbf{x}'||^2\right] \text{ is a constant} \\
&= C \sum_{n=0}^{\infty} \frac{\langle \mathbf{x}, \mathbf{x}' \rangle^n}{n!} && \text{Taylor expansion of } e^x \\
&= C \sum_{n=0}^{\infty} \frac{K_{\mathrm{poly(n)}}(\mathbf{x}, \mathbf{x}')}{n!}
\end{aligned}
$$

We see that the RBF kernel is formed by taking an infinite sum over polynomial kernels.

As proven previously, recall that the sum of two kernels

$$K_c(\mathbf{x}, \mathbf{x}') := K_a(\mathbf{x}, \mathbf{x}') + K_b(\mathbf{x}, \mathbf{x}')$$

implies that the ψ_c function is defined so that it forms vectors of the form

$$\psi_c(\mathbf{x}) := (\psi_a(\mathbf{x}), \psi_b(\mathbf{x}))$$

That is, the vector $\psi_c(\mathbf{x})$ is a tuple where the first element of the tuple is the vector $\psi_a(\mathbf{x})$ and the second element is $\psi_b(\mathbf{x})$. The inner-product on the vector space of ψ_c is defined as

$$\langle \psi_c(\mathbf{x}), \psi_c(\mathbf{x}') \rangle := \langle \psi_a(\mathbf{x}), \psi_a(\mathbf{x}') \rangle + \langle \psi_b(\mathbf{x}), \psi_b(\mathbf{x}') \rangle$$

For Euclidean vector spaces, this means that $\psi_c(\mathbf{x})$ is the vector formed by appending the elements of $\psi_b(\mathbf{x})$ onto the $\psi_a(\mathbf{x})$ and that

$$\begin{aligned} \langle \psi_c(\mathbf{x}), \psi_c(\mathbf{x}') \rangle &:= \sum_i^{\text{dimension}(a)} \psi_{a,i}(\mathbf{x})\psi_{a,i}(\mathbf{x}') + \sum_j^{\text{dimension}(b)} \psi_{b,j}(\mathbf{x})\psi_{b,j}(\mathbf{x}') \\ &= \sum_i^{\text{dimension}(a)+\text{dimension}(b)} \psi_{c,i}(\mathbf{x})\psi_{c,i}(\mathbf{x}') \end{aligned}$$

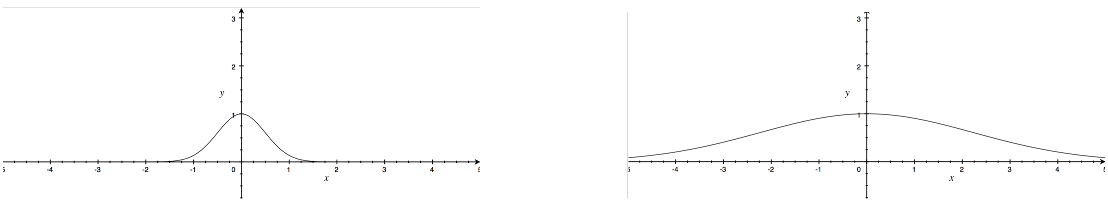
Since the RBF is an infinite sum over such appendages of vectors, we see that the projections is into a vector space with infinite dimension.

□

The γ parameter

Recall a kernel expresses a measure of similarity between vectors. The RBF kernel represents this similarity as a decaying function of the distance between the vectors (i.e. the squared-norm of their distance). That is, if the two vectors are close together then, $||\mathbf{x} - \mathbf{x}'||$ will be small. Then, so long as $\gamma > 0$, it follows that $-\gamma\,||\mathbf{x} - \mathbf{x}'||^2$ will be larger. Thus, closer vectors have a larger RBF kernel value than farther vectors. This function is of the form of a bell-shaped curve.

The γ parameter sets the width of the bell-shaped curve. The larger the value of γ the narrower will be the bell. Small values of γ yield wide bells. This is illustrated in Figure 1.



(a) **(b)**

Figure 1: (a) Large γ. (b) Small γ.

New See Our Resualt

The DataSet : wine

Classes : 3
Samples per class : [59,71,48]
Samples total : 178
Dimensionality : 13
Features : real, positive

---

without using Rbfspace and just use center of data:

we use distance sqeuclidean

and new we use RBFspace with different gamma:

when we use gamma = 1 , our accuracy decrease :(so we change gamma to 3

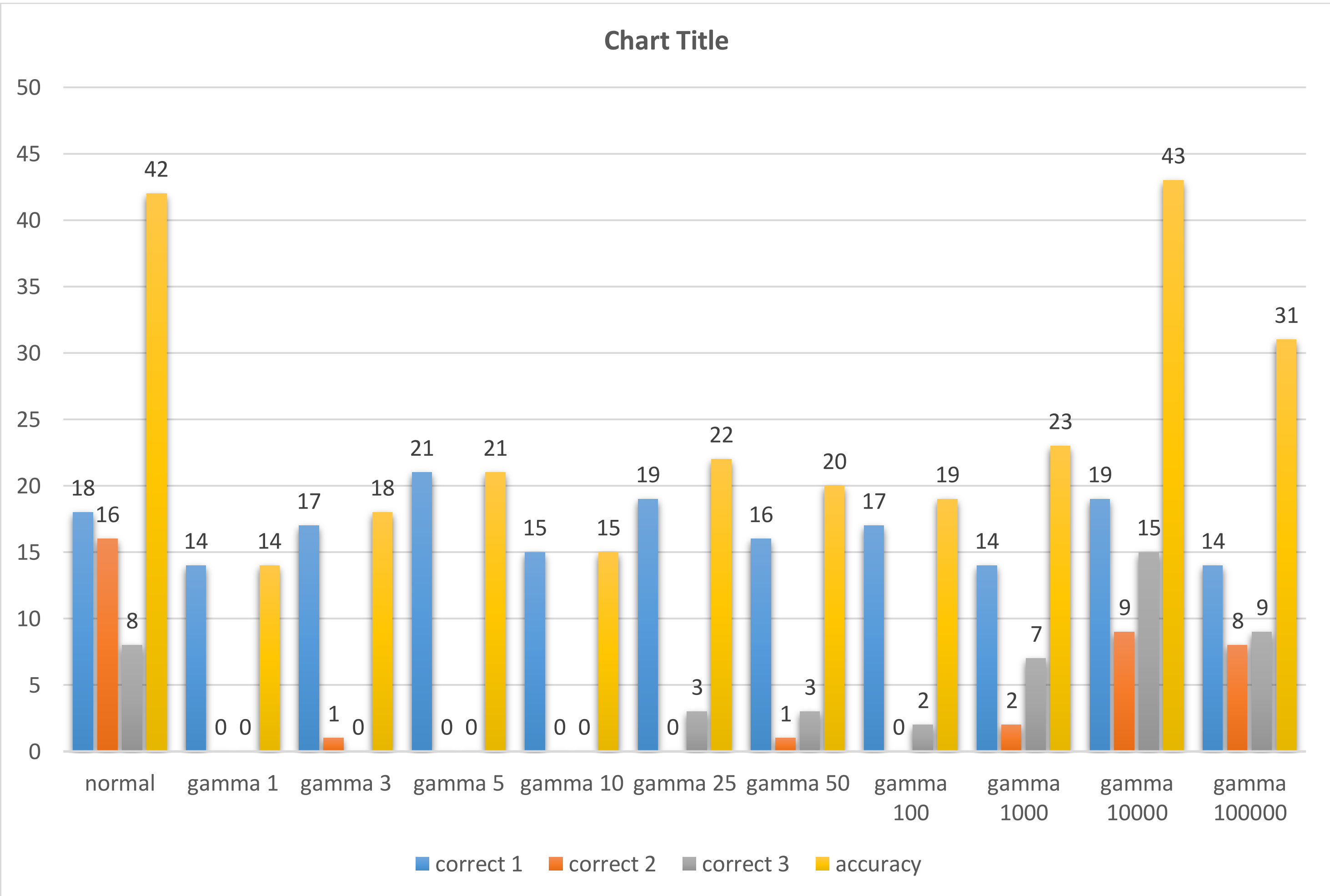
a little better but it still smaller than normal model

so we change gamma to 5

this gamma good for group 1 but not good for group 2 and 3 :(

we change gamma to 100 but Because of we use 1 gamma for all RBF space with increase the data fit with group 2

the increase more gamma make data fit with group 3 and 2 more



Number of data is 54

The result

1. It’s better to use different gamma for each group:

   We see gamma = 5 for group 1 and gamma=1000 for group 2 and gamma=100000 for group 3 can get better result . we get accuracy = 89

2. It’s better to use local search algorithm like Hill climbing , Simulated annealing (suited for either local or global search) ,Tabu search and Late acceptance hill climbing

Thank you :)