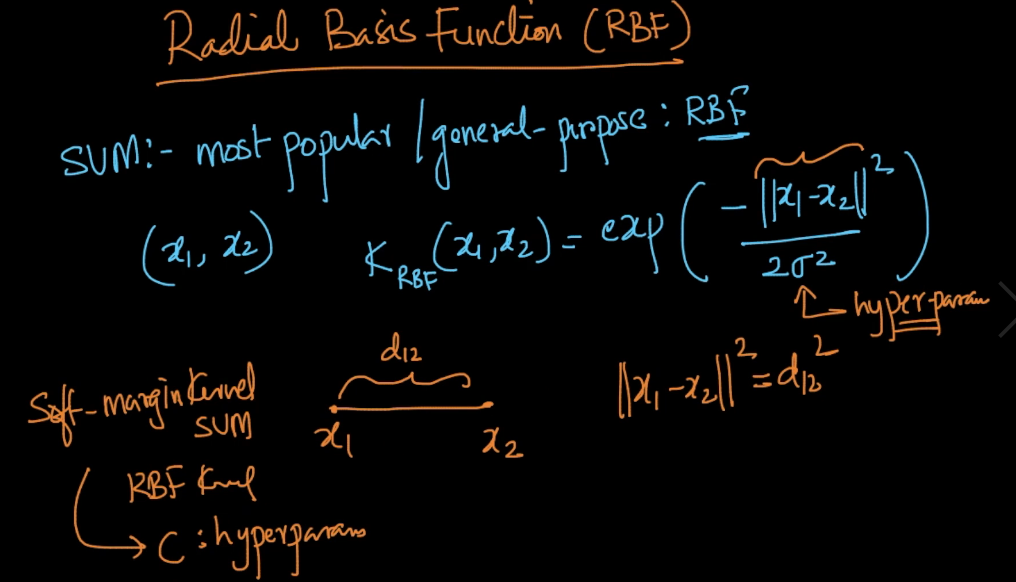
**RBF-Kernel**

RBF is a most popular and general purpose kernel function.

RBF kernel is represented as shown in below image.

In this ||x1 – x2||2 is a distance between points x1 and x2 i.e d122

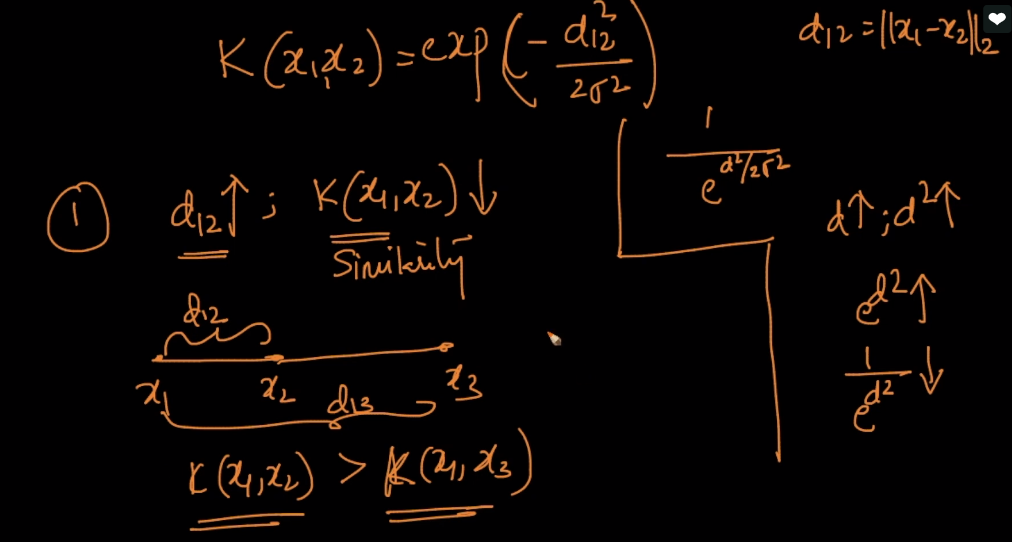
And sigma is hyperparameter.



As kernel = e(-x) therefore if x increases kernel decreases and vice versa.

Similarly in our equation as distance increases, kernel/similarity decreases.

In below fig. as x1 and x2 have less distance than x1 and x3 therefore x1 and x2 are more similar than x1 and x3



Now we check what happen when we change hyperparameter sigma.

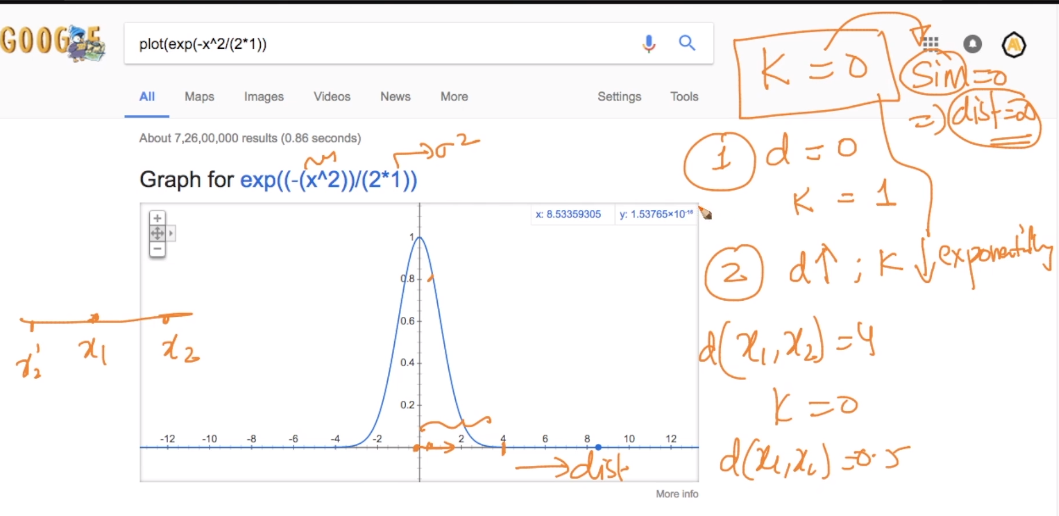
So first we take sigma = 1 .

In below fig. x-axis is distance and y-axis is sigma.

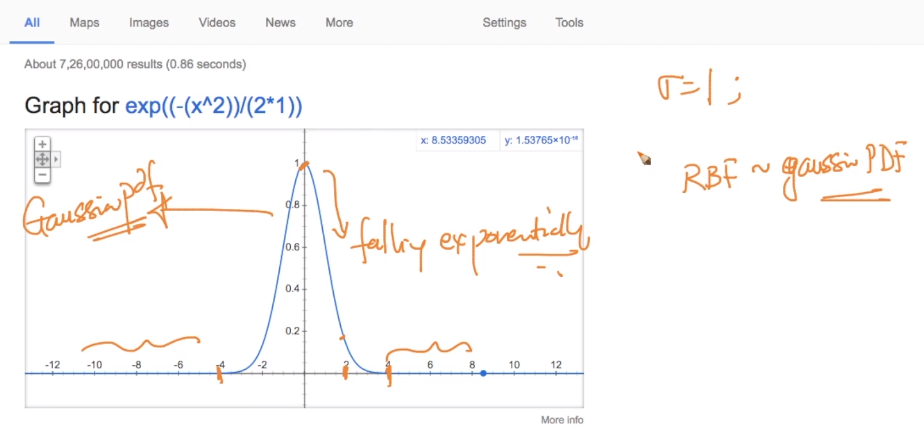
If d = 0 then k=1

And as we increase d(on any side +ve or -ve because we are taking square of distance), kernel decreases exponentialy.

So if d = 4 then k = 0 i.e there is 0 similarity b/w this two points

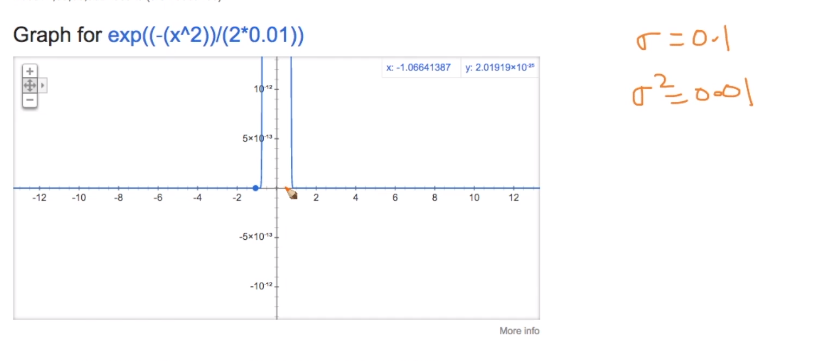


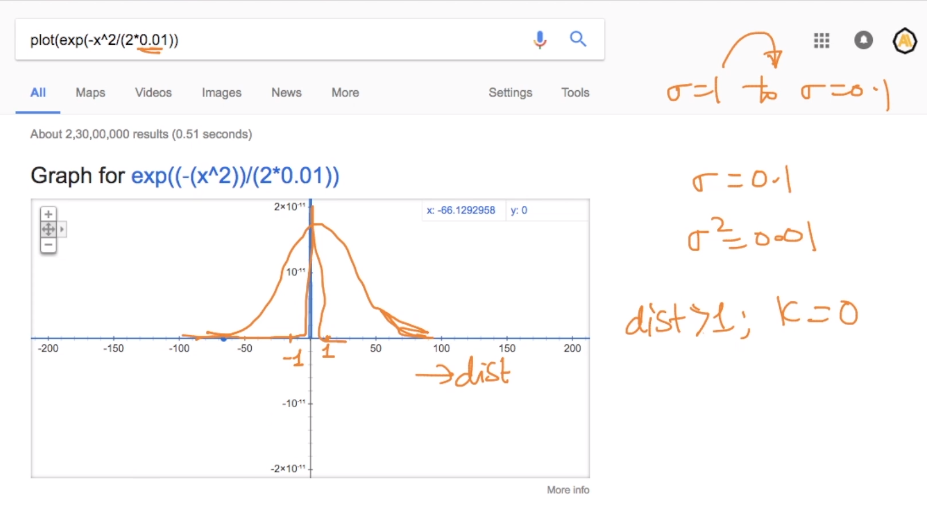
At sigma =1 it is almost same as gaussian pdf.



Below image shows kernel where sigma = 0.1 i.e sigma2 = 0.01

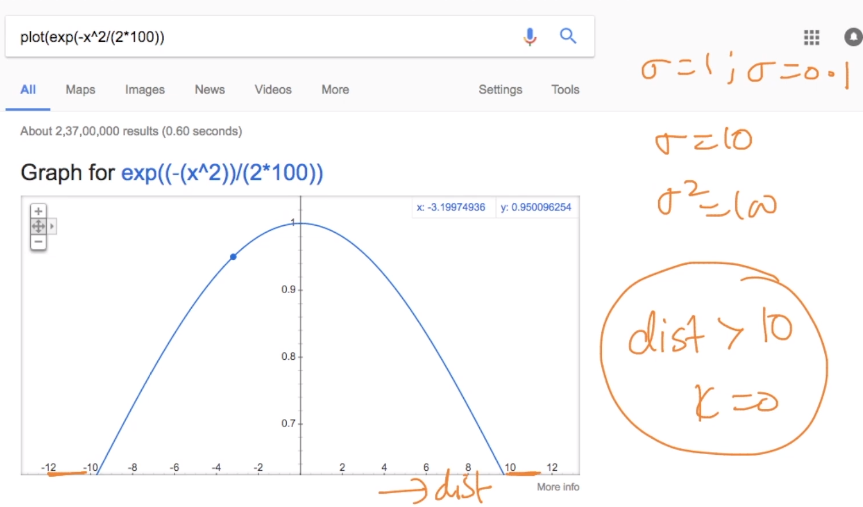
For this kernel if distance > 1 then k = 0





Now we take sigma = 10 i.e sigma2 = 100

In this curve becomes broader so in this if d>10 then only k = 0

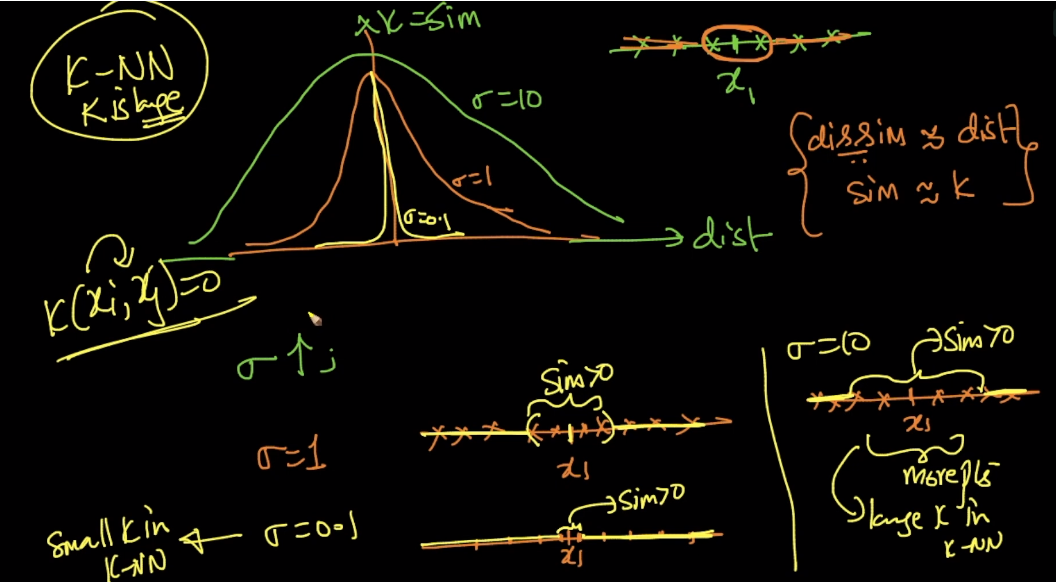


Below image shows the curve of kernels for different sigma and also when sigma increases similarity window also increases.

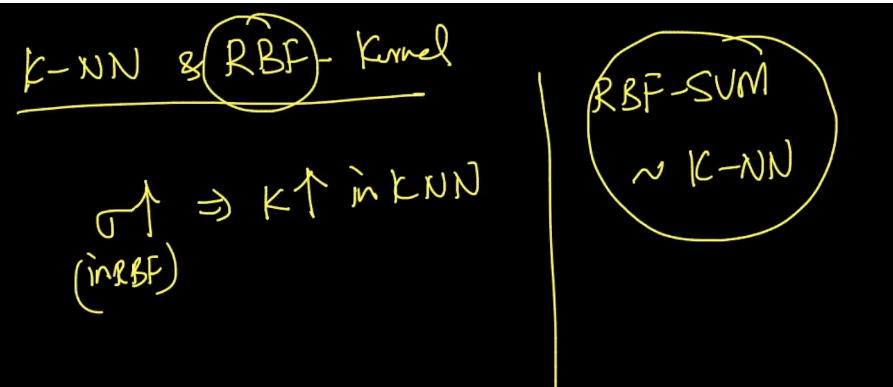
i.e if sigma = 0.1 then two points to be very close to be similar otherwise similarity becomes 0

if sigma = 1 then points need to be close but not as in 0.1 i.e similarity window increases.

If sigma =1 0 in this similarity window more increases.

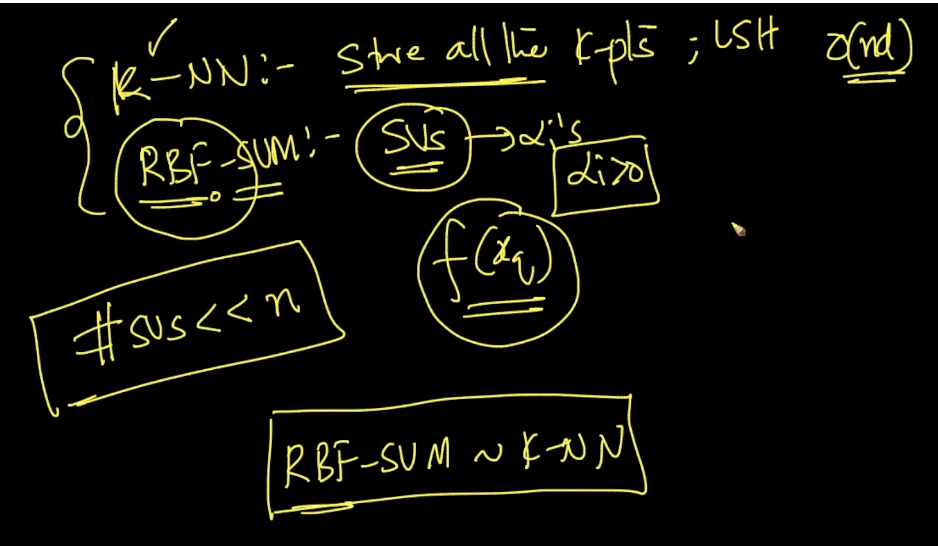


RBF kernel and k-NN are related because as hyperparameter sigma in RBF increases near by points in similarity window increases just like as k increases k-nn near by points also increases.



As k-nn and RBF svm similar they are both general purpose algorithm. But k-nn is slow it’s time complexity is more.

But SVM is fast because if got alphai’s of support vector because alpha of non-support vector is zero then we can find output of xq by passing it in f(xq)

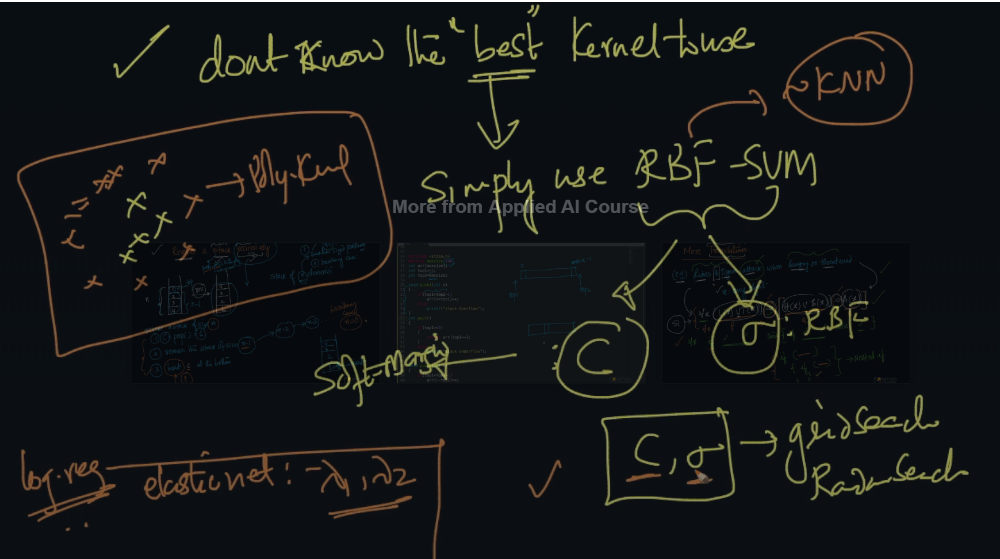


If we don’t know the best kernel tune then simply use RBF-SVM.

RBF-SVM have two hyperparameter c : (soft-margin) and sigma : (RBF)

We have to find this values of hyperparameter by grid search and random search.

Just like we do in elastic net l1 and l2 of logistic regression.



Comments :

