

# Radial basis function kernel

From Wikipedia, the free encyclopedia

In machine learning, the **(Gaussian) radial basis function kernel**, or **RBF kernel**, is a popular kernel function used in various kernelized learning algorithms. In particular, it is commonly used in support vector machine classification.<sup>[1]</sup>

The RBF kernel on two samples  $\mathbf{x}$  and  $\mathbf{x}'$ , represented as feature vectors in some *input space*, is defined as<sup>[2]</sup>

$$K(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2}\right)$$

$\|\mathbf{x} - \mathbf{x}'\|^2$  may be recognized as the squared Euclidean distance between the two feature vectors.  $\sigma$  is a free parameter. An equivalent, but simpler, definition involves a parameter  $\gamma = \frac{1}{2\sigma^2}$ :

$$K(\mathbf{x}, \mathbf{x}') = \exp(-\gamma\|\mathbf{x} - \mathbf{x}'\|^2)$$

Since the value of the RBF kernel decreases with distance and ranges between zero (in the limit) and one (when  $\mathbf{x} = \mathbf{x}'$ ), it has a ready interpretation as a similarity measure.<sup>[2]</sup> The feature space of the kernel has an infinite number of dimensions; for  $\sigma = 1$ , its expansion is:<sup>[3]</sup>

$$\begin{aligned}\exp\left(-\frac{1}{2}\|\mathbf{x} - \mathbf{x}'\|^2\right) &= \sum_{j=0}^{\infty} \frac{(\mathbf{x}^\top \mathbf{x}')^j}{j!} \exp\left(-\frac{1}{2}\|\mathbf{x}\|^2\right) \exp\left(-\frac{1}{2}\|\mathbf{x}'\|^2\right) \\ &= \sum_{j=0}^{\infty} \sum_{\sum n_i=j} \exp\left(-\frac{1}{2}\|\mathbf{x}\|^2\right) \frac{x_1^{n_1} \cdots x_k^{n_k}}{\sqrt{n_1! \cdots n_k!}} \exp\left(-\frac{1}{2}\|\mathbf{x}'\|^2\right) \frac{x_1'^{n_1} \cdots x_k'^{n_k}}{\sqrt{n_1! \cdots n_k!}}\end{aligned}$$

## Contents

- 1 Approximations
- 2 External links
- 3 See also
- 4 References

# Approximations

Because support vector machines and other models employing the kernel trick do not scale well to large numbers of training samples or large numbers of features in the input space, several approximations to the RBF kernel (and similar kernels) have been introduced.<sup>[4]</sup> Typically, these take the form of a function  $z$  that maps a single vector to a vector of higher dimensionality, approximating the kernel:

$$\langle z(\mathbf{x}), z(\mathbf{x}') \rangle \approx \langle \varphi(\mathbf{x}), \varphi(\mathbf{x}') \rangle = K(\mathbf{x}, \mathbf{x}')$$

where  $\varphi$  is the implicit mapping embedded in the RBF kernel.

One way to construct such a  $z$  is to randomly sample from the Fourier transformation of the kernel.<sup>[5]</sup> Another approach uses the Nyström method to approximate the eigendecomposition of the Gram matrix  $K$ , using only a random sample of the training set.<sup>[6]</sup>


## External links

- Kernels Part 1: What is an RBF Kernel? Really? ([http://charlesmartin14.wordpress.com/2012/02/06/kernels\\_part\\_1/](http://charlesmartin14.wordpress.com/2012/02/06/kernels_part_1/))

## See also

- Gaussian function
- Kernel (statistics)
- Polynomial kernel
- Radial basis function
- Radial basis function network

## References

1. Yin-Wen Chang, Cho-Jui Hsieh, Kai-Wei Chang, Michael Ringgaard and Chih-Jen Lin (2010). "Training and testing low-degree polynomial data mappings via linear SVM" (<http://jmlr.org/papers/v11/chang10a.html>). *J. Machine Learning Research* **11**:1471–1490.
2. Jean-Philippe Vert, Koji Tsuda, and Bernhard Schölkopf (2004). "A primer on kernel methods". (<http://cbio.enscm.fr/~jvert/publi/04kmcbbook/kernelprimer.pdf>) *Kernel Methods in Computational Biology*.
3. Shashua, Amnon (2009). "Introduction to Machine Learning: Class Notes 67577". arXiv:0904.3664v1 (<https://arxiv.org/abs/0904.3664v1>)  [cs.LG (<https://arxiv.org/archive/cs.LG>)].
4. Andreas Müller (2012). Kernel Approximations for Efficient SVMs (and other feature extraction methods) (<http://peekaboo-vision.blogspot.de/2012/12/kernel-approximations-for-efficient.html>).
5. Ali Rahimi and Benjamin Recht (2007). "Random features for large-scale kernel machines" (<http://www.eecs.berkeley.edu/~brecht/papers/07.rah.rec.nips.pdf>). *Neural Information Processing Systems*.
6. C.K.I. Williams and M. Seeger (2001). "Using the Nyström method to speed up kernel machines" (<http://papers.nips.cc/paper/1866-using-the-nystrom-method-to-speed-up-kernel-machines>). *Advances in Neural Information Processing Systems*.

Retrieved from "[https://en.wikipedia.org/w/index.php?title=Radial\\_basis\\_function\\_kernel&oldid=777638273](https://en.wikipedia.org/w/index.php?title=Radial_basis_function_kernel&oldid=777638273)"

Categories: Kernel methods for machine learning

---

- This page was last edited on 28 April 2017, at 11:15.
- Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.