SVM notes

Justin

# Optimum number of supporting vectors

If the data is linearly separable, a smaller number of supporting vectors may be desirable, and vice versa.

Check this [link](https://github.com/rasbt/python-machine-learning-book/blob/master/faq/num-support-vectors.md) for details.

# Kernel selection

The rule of thumb is:  use linear SVMs (or logistic regression) for linear problems, and nonlinear kernels such as the Radial Basis Function kernel for non-linear problems.

In any case, don’t bother too much about the polynomial kernel. In practice, it is less useful for efficiency (computational as well as predictive) performance reasons.

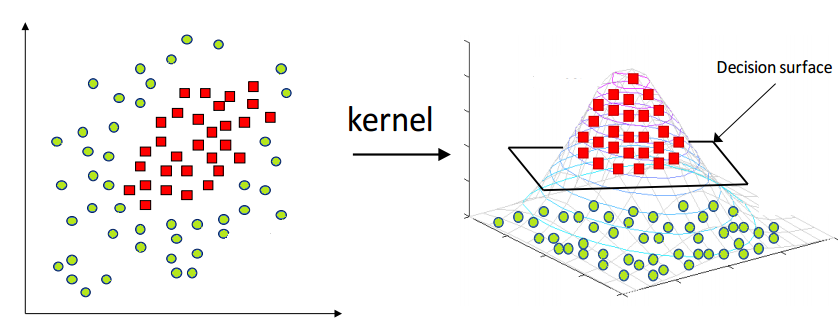


Figure . The kernel magic projects the data to higher dimensions and hence makes it linearly separable

# Activation function and loss function