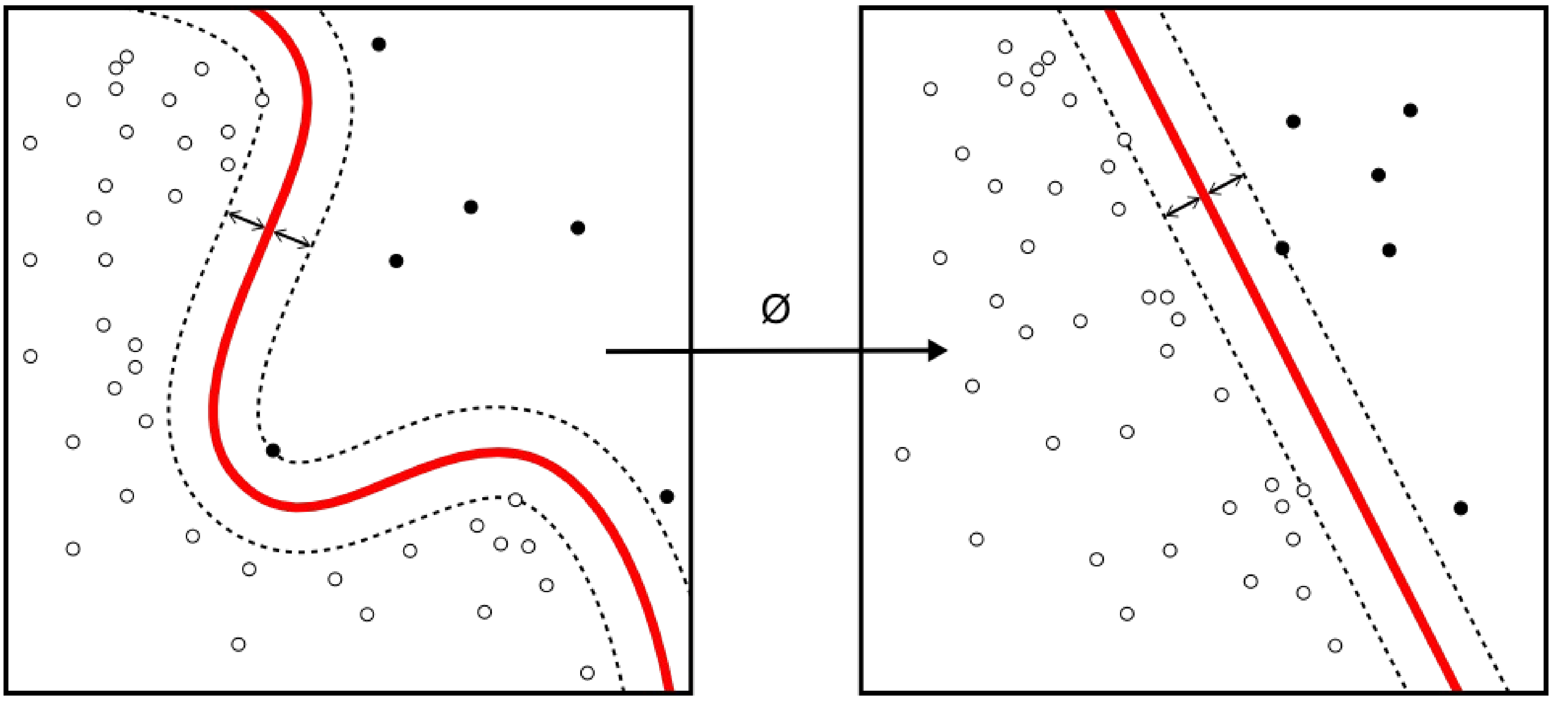
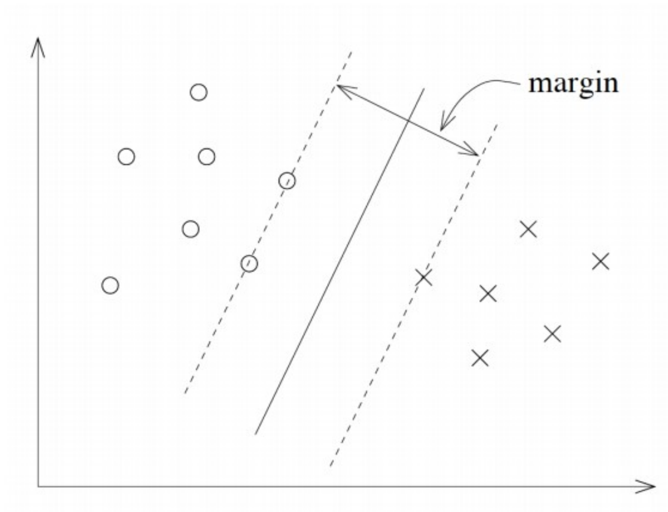
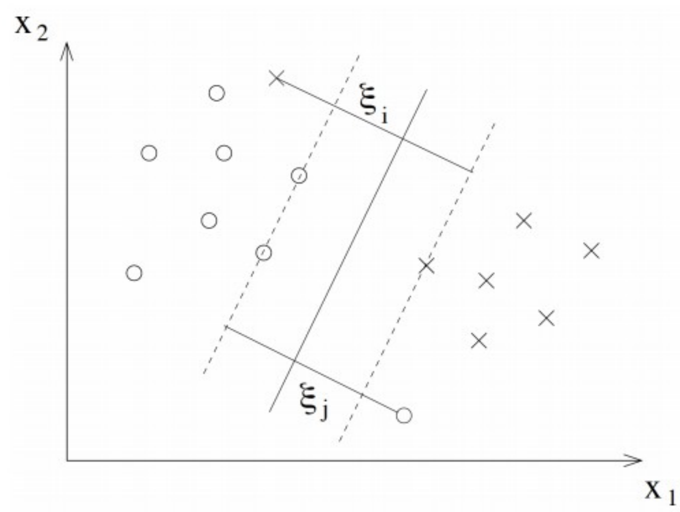
SVM



Hyperparameter C





- In this general case, the problem of SVM can be stated as:
  - Find the hyperplane (w,b) that minimizes

$$\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i$$

  - Under the condition that, for 1 ≤ i ≤ N:

$$y_i(w' x_i + b) \geq 1 - \xi_i$$

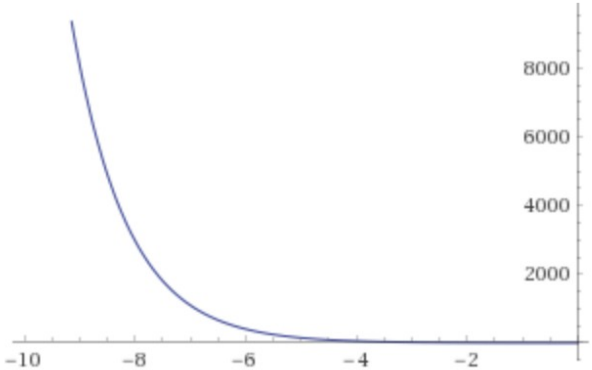
rbf Kernel & gamma

$$w'x + b = \sum_{i=1}^{n} w_i x_i + b = 0$$

$$w\varphi(x) + b = \left(\sum_{i=1}^{N} \alpha_i y_i \varphi(x_i)\right)\varphi(x) + b$$

$$= \sum_{i=1}^{N} \alpha_i y_i \langle \varphi(x_i), \varphi(x) \rangle + b$$

$$= \sum_{i=1}^{N} \alpha_i y_i \kappa(x_i, x) + b$$

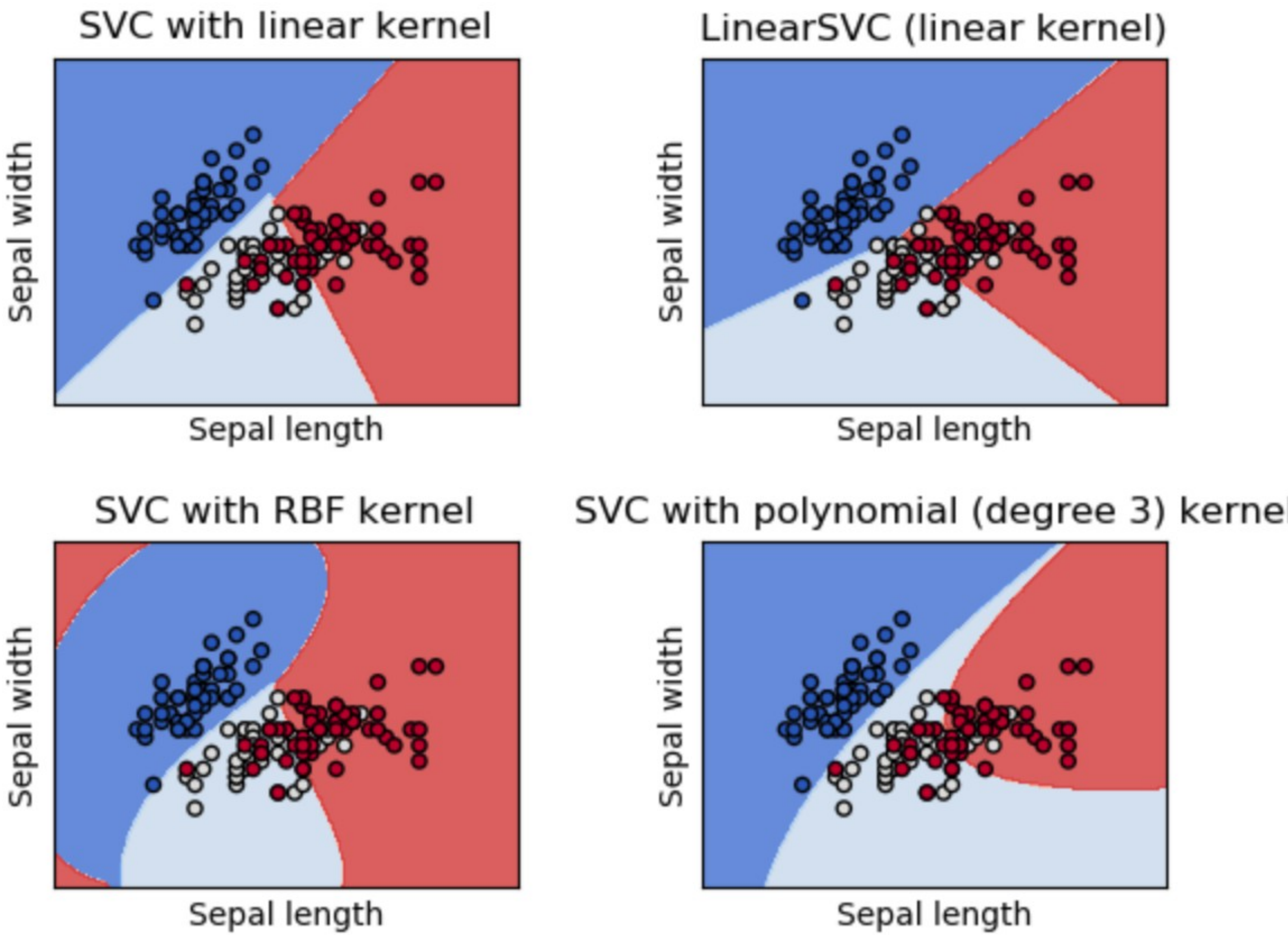


$$\boxed{\kappa(x,y) = \langle x, y \rangle \text{ linear kernel}}$$

$$\kappa(x,y) = \left(\gamma \langle x, y \rangle + r\right)^d, \gamma > 0 \text{ polynomial kernel}$$

$$\boxed{\kappa(x,y) = \exp\left(-\gamma \|x - y\|^2\right), \gamma > 0 \text{ RBF (Gaussian) kernel}}$$

$$\kappa(x,y) = \tanh\left(\gamma \langle x, y \rangle + r\right) \text{ sigmoid kernel}$$


SVC with linear kernel
Sepal width
Sepal length
LinearSVC (linear kernel)
Sepal width
Sepal length
SVC with RBF kernel
Sepal width
Sepal length
SVC with polynomial (degree 3) kernel
Sepal width
Sepal length

Setup

```
import numpy as np
from sklearn.svm import SVC

train = np.genfromtxt('../data/MNIST/train_med.csv', delimiter=',')
test = np.genfromtxt('../data/MNIST/test.csv', delimiter=',')

y = train[:60000, 0]
X = train[:60000, 1:] / 255.

y_test = test[:, 0]
X_test = test[:, 1:] / 255.
```

Calculation

```
accuracies = np.zeros((6, 6))

for i in range(0, 6):
    for j in range(0, 6):
                                        'rbf'
        svm = SVC(verbose=False, kernel='linear',
                  cache_size=4000, gamma=2**(j-5), C=10**(i+5))
        svm.fit(X, y)
        pred = svm.predict(X_test)
        accuracies[i,j] = ((y_test == pred).sum() + 0.0) / len(pred)
        print ('Gamma={0}, C={1}, Acc={2}'
               .format(2**(j-5), 10**(i+5), accuracies[i,j]))

print accuracies
```

Results

```
C Gamma Score Time
100000.0 0.03125 0.981745 -31.6
100000.0 0.03125 0.985336 -31.7
100000.0 0.03125 0.982500 -31.7
100000.0 0.03125 0.985090 -31.8
100000.0 0.03125 0.984328 -32.2
100000.0 0.0625 0.979753 -75.6
100000.0 0.0625 0.978250 -76.0
100000.0 0.0625 0.980841 -76.3
100000.0 0.0625 0.976494 -74.4
100000.0 0.0625 0.980077 -75.2
100000.0 0.125 0.892795 -188.8
100000.0 0.125 0.879853 -188.8
100000.0 0.125 0.888000 -189.3
1000000.0 0.03125 0.985090 -30.3
100000.0 0.125 0.883721 -188.2
100000.0 0.125 0.900717 -189.0
```