1.

"Soft" margin solution

The optimization problem becomes

$$\min_{\mathbf{w}\in\mathbb{R}^d, \xi_i\in\mathbb{R}^+} ||\mathbf{w}||^2 + C\sum_i^N \xi_i$$

subject to

$$y_i\left(\mathbf{w}^\top \mathbf{x}_i + b\right) \geq 1 - \xi_i \text{ for } i = 1 \dots N$$

For following formulation of SVM for C→ infinity, ε_i tends to what value ? What special case of SVM will we get with C → infinity ?

2.

Can α be zero for support vectors? Support your answer with proper reasoning.

Special transformations

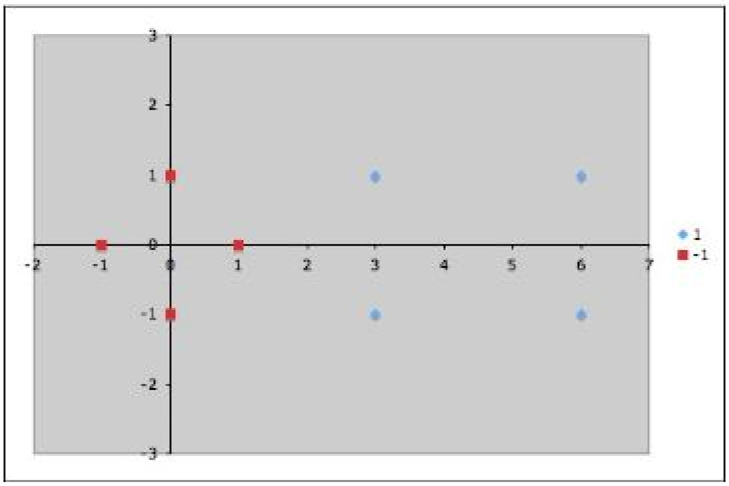
$$\Phi : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to \begin{pmatrix} x_1^2 \\ x_2^2 \\ \sqrt{2}x_1x_2 \end{pmatrix} \quad \mathbb{R}^2 \to \mathbb{R}^3$$

$$\begin{aligned} \Phi(\mathbf{x})^\top\Phi(\mathbf{z}) &= \left(x_1^2, x_2^2, \sqrt{2}x_1x_2\right)\begin{pmatrix} z_1^2 \\ z_2^2 \\ \sqrt{2}z_1z_2 \end{pmatrix} \\ &= x_1^2z_1^2 + x_2^2z_2^2 + 2x_1x_2z_1z_2 \\ &= (x_1z_1 + x_2z_2)^2 \\ &= (\mathbf{x}^\top\mathbf{z})^2 \end{aligned}$$

3.

Above transformation of quantity $\Phi(\mathbf{x})^\top\Phi(\mathbf{z})$ to $(\mathbf{x}^\top\mathbf{z})^2$ to use in deriving SVM parameters is called as kernel trick. Instead of $(x^Tz)^2$ let's say you wanted to convert to $(x^Tz)^3$. What would $\Phi : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$? That is define mapping similar to above derivation which finally gives cube of final result instead of square.

4.



In the above figure,Blue diamonds are the positively labelled data points and red squares denote negatively labelled data points in a 2-D plane. By inspection, it should be obvious that there are three support vectors:

s1=(1,0)
s2=(3,1)
s3=(3,-1)

Find the discriminating hyperplane that separates the positive from the negative examples?

Hint: Since the data is linearly separable, we can use a linear SVM.(that is, one whose mapping function Φ() is the identity function).