

## Multimedia Appendix 2: SVM

The SVM computation process with mathematical formulas is explained as follows.

SVMs try to maximize the margin between classes (here using the simple linear feature space  $x_i \cdot x_j$ ), by finding the optimal  $\alpha_i$  values in the following quadratic programming problem (represented in dual Lagrangian form where  $C$  is a constant that bounds the misclassification error):

$$\max \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (2.1)$$

subject to:  $0 \leq \alpha_i \leq C$  and  $\sum_{i=1}^N \alpha_i y_i = 0$

Unlabelled instances are classified using the learned parameters  $\alpha_i$  and bias  $b$ , by taking the sign of the following decision function:

$$f(x) = \sum_{i=1}^N \alpha_i y_i (x \cdot x_i) + b \quad (2.2)$$