**Machine Learning Worksheet 05**

# Linear Classification

## 1 Linear separability

**Problem 1:** Given a set of data points $\mathcal{X} = \{\boldsymbol{x}_i\}_{i=1}^N$, we can define the *convex hull* $\mathrm{co}\mathcal{X}$ to be the set of all points $\boldsymbol{x}$ given by

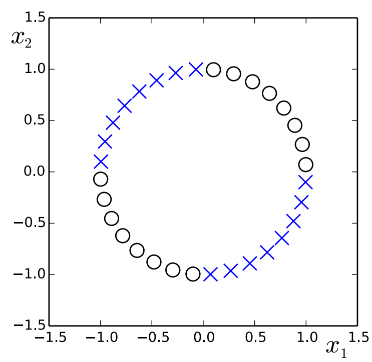
$$\mathrm{co}\mathcal{X} = \{\boldsymbol{x} : \boldsymbol{x} = \sum_i \alpha_i \boldsymbol{x}_i, \alpha_i \geq 0, \sum_i \alpha_i = 1\}$$

Consider a second set of points $\mathcal{Y} = \{\boldsymbol{y}_j\}_{j=1}^M$ together with its corresponding convex hull. By definition, the two sets of points will be linearly separable if there exists a vector $\boldsymbol{w}$ and a scalar $w_0$ such that $\boldsymbol{w}^T\boldsymbol{x}_i + w_0 > 0$ for all $\boldsymbol{x}_i \in \mathcal{X}$, and $\boldsymbol{w}^T\boldsymbol{y}_j + w_0 < 0$ for all $\boldsymbol{y}_j \in \mathcal{Y}$. Show that if their convex hulls intersect, the two sets of points cannot be linearly separable.

**Problem 2:** Show that for a linearly separable data set, the maximum likelihood solution for the logistic regression model is obtained by finding a vector $\boldsymbol{w}$ whose decision boundary $\boldsymbol{w}^T\boldsymbol{x} = 0$ separates the classes and then taking the magnitude of $\boldsymbol{w}$ to infinity. Assume that $\boldsymbol{w}$ contains the bias term.

How can we prevent this?

**Problem 3:** Which basis function $\phi(x_1, x_2)$ makes the data in the example below linearly separable (crosses in one class, circles in the other)?




## 2 Basis functions

**Problem 4:** The decision boundary for a linear classifier on two-dimensional data crosses axis $x_1$ at 2 and $x_2$ at 5. Write down the general form of this linear classifier model with a bias term (how many parameters do you need, given the dimensions?) and calculate possible coefficients (parameters).