# COMP9417 – Machine Learning

# Tutorial: Kernel Methods

# z5325987

## Question 4. (Kernels and their Feature Representations)

(a) $k(x,y) = (2 < x,y > +3)^3 = (2\sum_{i=1}^{n} x_i y_i + 3)^3 = 8(\sum_{i=1}^{n} x_i y_i)^3 + 54\sum_{i=1}^{n} x_i y_i + 36(\sum_{i=1}^{n} x_i y_i)^2 + 27$

$$= 8\sum_{i=1}^{n}(x_i)^3(y_i)^3 + 24\sum_{i=2}^{n}\sum_{j=1}^{i-1}\left(x_i^2 x_j\right)\left(y_i^2 y_j\right) + 54\sum_{i=1}^{n} x_i y_i + 36\sum_{i=1}^{n}(x_i)^2(y_i)^2$$

$$+72\sum_{i=2}^{n}\sum_{j=1}^{i-1}\left(x_i x_j\right)\left(y_i y_j\right) + 27$$

Therefore, $\varphi(x) =< 8x_n^3, \dots, 8x_1^3, 2\sqrt{6}x_n^2 x_{n-1}, \dots, 2\sqrt{6}x_n^2 x_1, 2\sqrt{6}x_{n-1}^2 x_{n-2}, \dots, 2\sqrt{6}x_{n-1}^2 x_1, \dots, 2\sqrt{6}x_2^2 x_1, \dots, 54x_n, \dots, 54x_1, 36x_n^2, \dots, 36x_1^2, 6\sqrt{2}x_n x_{n-1}, \dots, 6\sqrt{2}x_n x_1, 6\sqrt{2}x_{n-1}x_{n-2}, \dots, 6\sqrt{2}x_{n-1}x_1, \dots, 6\sqrt{2}x_2 x_1, 3\sqrt{3} >$

## Question 7. (Support Vector Machines)

1. The Gram matrix is

$$X^T X = \begin{bmatrix} 10 & 5 & 3 \\ 5 & 5 & 1 \\ 3 & 1 & 1 \end{bmatrix} \quad X'^T X' = \begin{bmatrix} 10 & 5 & -3 \\ 5 & 5 & -1 \\ -3 & -1 & 1 \end{bmatrix}$$

2. Expression to be minimized

$$L(w,t,\alpha) = \frac{1}{2}\left||w|\right|^2 - \sum_{i=1}^{N} \alpha_i y_i (w x_i + t) + \sum_{i=1}^{N} \alpha_i$$

3 and 4. Take partial derivatives and set 0

$$\nabla_w L(w,t,\alpha) = w - \sum_{i=1}^{N} \alpha_i y_i x_i = 0$$

$$\nabla_t L(w,t,\alpha) = -\sum_{i=1}^{N} \alpha_i y_i = 0$$

5. Solve for w

$$w = \sum_{i=1}^{N} \alpha_i y_i x_i$$

6. Solve for t

$$t = y_j - \sum_{\mathrm{i=1}}^{\mathrm{N}} \alpha_i y_i (x_i x_j)$$