The Exercise of SVM

Yanwei Fu

October 19, 2016

1 Kernel

(a) Determine the mapping $\phi(x)$ such that the kernel

$$k(\mathbf{x}, \mathbf{z}) = \left(c + \mathbf{x}^T \mathbf{z}\right)^2 = \phi(\mathbf{x})\,\phi(\mathbf{z})$$

where $\mathbf{x} = (x_1, x_2)^T$ and $\mathbf{z} = (z_1, z_2)^T$.

(b) Show, by a sketch, that an XOR is not linearly separable, but that after the mapping $\phi(x)$ with $c = 0$ it is linearly separable.

2 Steepest Gradient Descendent of SVM

Show that if the SVM cost function is written as

$$\mathcal{C}(\mathbf{w}) = \frac{1}{N}\Sigma_{i=1}^{N}\left(\frac{\lambda}{2} \parallel \mathbf{w} \parallel^2 + \max(0, 1 - y_i f(\mathbf{x}_i))\right)$$

where $f(\mathbf{x}_i) = \mathbf{w}^T \mathbf{x}_i$, then using using steepest descent optimization, $\mathbf{w}_{t+1}$ may be learnt from $\mathbf{w}_t$ by cycling through the data with the following update rule,

$$\begin{aligned} \mathbf{w}_{t+1} &\leftarrow (1 - \eta\lambda)\,\mathbf{w}_t + \eta y_i \mathbf{x}_i \; if \; y_i \mathbf{w}^T \mathbf{x}_i < 1 \\ &\leftarrow (1 - \eta\lambda)\,\mathbf{w}_t \quad otherwise \end{aligned}$$

where η is the learning rate.