Exercise Sheet 10

Machine Learning 2, SS16

July 11, 2016

Mario Tambos, 380599; Viktor Jeney, 348969; Sascha Huk, 321249; Jan Tinapp, 0380549

Exercise 1a

For arbitrary n let $\vec{x}_1, \dots, \vec{x}_n \in \mathcal{X}$ be arbitrary samples and $\vec{v} \in \mathbb{R}^n$ be an arbitrary vector.

$$\forall_{l=1}^{L}. \sum_{i,j=1}^{n} v_i v_j k_l(x_i, x_j) \geq 0 \Longrightarrow \sum_{l=1}^{L} \sum_{i,j=1}^{n} v_i v_j k_l(x_i, x_j) \geq 0 \Longrightarrow \sum_{i,j=1}^{n} v_i v_j \sum_{l=1}^{L} k_l(x_i, x_j) \geq 0$$

$$\overset{\vec{\beta} \geq 0}{\Longrightarrow} \sum_{i,j=1}^{n} v_i v_j \sum_{l=1}^{L} \beta_l k_l(x_i, x_j) \geq 0 \overset{\text{Def. k}}{\Longrightarrow} \sum_{i,j=1}^{n} v_i v_j k(x_i, x_j) \geq 0$$

Exercise 1b

Let $x, x' \in \mathcal{X}$ be two arbitrary samples.

$$\sum_{l=1}^{L} \beta_l k_l(x, x') = \sum_{l=1}^{L} \beta_l \phi_l(x)^\top \phi_l(x') = \sum_{l=1}^{L} (\sqrt{\beta_l} \phi_l(x)^\top)(\sqrt{\beta_l} \phi_l(x'))$$

$$= \underbrace{\begin{bmatrix} \sqrt{\beta_1}\phi_1(x)^\top & \dots & \sqrt{\beta_L}\phi_L(x)^\top \end{bmatrix}}_{\phi(x)^\top} \underbrace{\begin{bmatrix} \sqrt{\beta_1}\phi_1(x') \\ \dots \\ \sqrt{\beta_L}\phi_L(x') \end{bmatrix}}_{\phi(x')}$$

$\phi(x)^\top$ and $\phi(x')$ are block partitioned matrices, i.e. the $\phi_l's$ are simply concatenated together in one very long vector. So, the result is:

$$\phi(x) = \begin{bmatrix} \sqrt{\beta_1}\phi_1(x) \\ \dots \\ \sqrt{\beta_L}\phi_L(x) \end{bmatrix}$$

Exercise 2a

We also encode the classes via the canonical base vectors $\mathbf{e}_y$ for all $y \in \{1, \dots, C\}$. For arbitrary n let $\vec{x}_1, \dots, \vec{x}_n \in \mathcal{X}$ be arbitrary samples and $\vec{v} \in \mathbb{R}^n$ be an arbitrary vector.

$$\sum_{i,j=1}^{n} v_i v_j k(x_i, x_j) \mathbb{1}_{[y_i = y_j]} = \sum_{i,j=1}^{n} v_i v_j k(x_i, x_j) \mathbf{e}_{y_i}^\top \mathbf{e}_{y_j}$$

$$= \sum_{i,j} v_i v_j \phi(x_i)^\top \phi(x_j) \mathbf{e}_{y_i}^\top \mathbf{e}_{y_j} = \sum_{i,j} v_i v_j \sum_k \phi(x_i)_k \phi(x_j)_k \sum_l (\mathbf{e}_{y_i})_l (\mathbf{e}_{y_j})_l$$

$$= \sum_{k,l} \sum_{i,j} v_i v_j \phi(x_i)_k \phi(x_j)_k (\mathbf{e}_{y_i})_l (\mathbf{e}_{y_j})_l = \sum_{k,l} \underbrace{(\sum_i v_i \phi(x_i)_k (\mathbf{e}_{y_i})_l)(\sum_j v_j \phi(x_j)_k (\mathbf{e}_{y_j})_l)}_{\text{sums are equal}}$$

$$= \sum_{k,l} \underbrace{(\sum_i v_i \phi(x_i)_k (\mathbf{e}_{y_i})_l)^2}_{\geq 0} \geq 0$$

Exercise 2b

We also encode the classes via the canonical base vectors $\mathbf{e}_y$ for all $y \in \{1, \dots, C\}$.

$$k(x, x') \mathbb{1}_{[y = y']} = \phi(x)^\top \phi(x') \mathbf{e}_y^\top \mathbf{e}_{y'} = \sum_{i=1}^{h} \phi_i(x) \phi_i(x') \sum_{c=1}^{C} (\mathbf{e}_y)_c (\mathbf{e}_{y'})_c = \sum_{i=1}^{h} \sum_{c=1}^{C} \phi_i(x) \phi_i(x') (\mathbf{e}_y)_c (\mathbf{e}_{y'})_c$$

$$= \sum_{c=1}^{C} \sum_{i=1}^{h} \phi_i(x) (\mathbf{e}_y)_c \phi_i(x') (\mathbf{e}_{y'})_c = < \underbrace{\phi(x) \times \mathbf{e}_y}_{\phi_{\text{struct}}(x,y)}, \underbrace{\phi(x') \times \mathbf{e}_{y'}}_{\phi_{\text{struct}}(x',y')} >_F$$

$< \cdot, \cdot >_F$ is the Frobenius inner product. If one prefers the standard scalar product (to see $\phi_{struct}(x)$ as a vector) one could convert the matrices ϕ_{struct} into vectors, e.g. by concatenating the row- or column vectors. Then using the standard scalar product on these vectors would yield the same result.