Exercise Sheet 7

Machine Learning 2, SS16

June 14, 2016

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

Exercise 1 - Weighted Degree Kernels

(a)

The weighted degree kernel seeks to incorporate the assumption that identical substrings in the same position of the two gene sequences have some special significance, and that subsequences of different length are differently important (depending on the particular choice of the β_m parameters).

(b)

Define everything in matrix notation in the following way:

$$K = [k(x_i, x_j)]_{i=1,\ldots,K, j=1,\ldots,K}$$
$$I(m, n) = [I(u_{m,n}(x_i) = u_{m,n}(x_j))]_{i=1,\ldots,K, j=1,\ldots,K}$$

Then we have to show, that for every $\alpha \in \mathbb{R}^K$

$$
\begin{aligned}
0 &\leq \sum_{i,j=1}^{K} \alpha_i \alpha_j k(x_i, x_j) \\
&= \alpha^T K \alpha \\
&= \alpha^T \left(\sum_{m=1}^{M} \beta_m \sum_{n=1}^{N-m+1} I(m, n) \right) \alpha \\
&= \sum_{m=1}^{M} \beta_m \sum_{n=1}^{N-m+1} \alpha^T I(m, n) \alpha
\end{aligned}
$$

Because all the $\beta_i, i = 1, \ldots, M$ are nonnegative and arbitrary, the following condition is sufficient:

$$0 \leq \alpha^T I(n, m) \alpha.$$

for every $m = 1, .., M, n = 1, \ldots, N - m + 1$. Now look at the form of the matrix $I(n, m)$. The diagonal is 1, as well as the entries (k, l) and (l, k), if $u_{m,n}(x_k) = u_{m,n}(x_l)$. All other entries are zero.

So define the set of entries, where $u_{m,n}(x_k) = u_{m,n}(x_l)$ and $k < l$ as $S(m, n)$ and define the matrices

$$[A(k, l)]_{i=1,\ldots K, j=1,\ldots,K} := \begin{cases} 1 & , (i, j) = (k, l) \, or \, (i, j) = (l, k) \\ 0 & , otherwise \end{cases}$$

Then the matrix $I(m, n)$ can be decomposed as a sum of the identity matrix E and the previously defined matrices:

$$I(m, n) = E + \sum_{(k,l) \in S(m,n)} A(k, l),$$

which leads to

$$\alpha^T I(m,n)\alpha = \alpha^T E\alpha + \sum_{(k,l)\in S(m,n)} \alpha^T A(k,l)\alpha$$

$$= \alpha^T \alpha + \sum_{(k,l)\in S(m,n)} \alpha^T A(k,l)\alpha$$

$$= \sum_{i=1,\ldots,K} \alpha_i^2 + \sum_{(k,l)\in S(m,n)} 2\alpha_k\alpha_l$$

$$= \sum_{(k,l)\in S(m,n)} \left(\alpha_k^2 + 2\alpha_k\alpha_l + \alpha_l^2\right) + \sum_{i\in R} \alpha_i^2$$

$$= \sum_{(k,l)\in S(m,n)} (\alpha_k + \alpha_l)^2 + \sum_{i\in R} \alpha_i^2$$

where R is defined as the set of $i = 1,..,K$, for which (i,j) and (j,i) are never in $S(m,n)$ (for any choice of j).(So all indexes, that where taken into the first sum)
Now we have a sum of squares, which is nonegative. Hooray!
So to conclude, we have for every $\alpha \in \mathbb{R}^K$ and every $m = 1,...,M, n = 1,...,N-m+1$:

$$\alpha^T I(m,n)\alpha \geq 0,$$

which gives us

$$\sum_{i,j=1}^K \alpha_i\alpha_j k(x_i,x_j) = \alpha^T K\alpha \geq 0$$