Задание №2.2. Ядра

На лекции мы увидели, что, используя различные ядра $K(x,z) = \phi(x)^T\phi(x)$, мы можем неявным образом отображать исходные данные в пространства признаков более высоких размерностей, и, соответственно, обучать нашу модель (например, SVM) в этих пространствах. Один из способов построения ядер заключается в том, чтобы сначала определить трансформацию ϕ, а затем постараться вывести соответствующую ему функцию K.

Однако в данном упражнении нас будет интересовать процесс непосредственного построения ядер. В частности, представим, что у нас есть функция $K(x,z)$, которая, как мы считаем, является адекватной мерой близости образцов x и z в нашей задаче, и мы хотим использовать эту функцию в качестве ядра в SVM классификаторе. Мы помним, что для того, чтобы $K(x,z)$ было корректным ядром, оно должно соответствовать скалярному произведению значений, отображенных с помощью какой-то подходящей функции ϕ в некоторое пространство большей размерности. Теорема Мерсера говорит нам, что $K(x,z)$ является ядром (Мерсера) тогда и только тогда, когда для любого конечного множества $\{x^{(1)},\dots,x^{(n)}\}$ квадратная матрица $K \in \mathbb{R}^{n\times n}$, задаваемая значениями $K_{ij} = K\left(x^{(i)},x^{(j)}\right)$, является симметричной положительной полуопределенной.

А теперь переходим к вопросу. Пусть:

- K_1 и K_2 являются ядрами на $\mathbb{R}^d \times \mathbb{R}^d$,
- $a \in \mathbb{R}^+$,
- $f: \mathbb{R}^d \to \mathbb{R}$,
- $\phi: \mathbb{R}^d \to \mathbb{R}^p$,
- K_3 – ядро на $\mathbb{R}^p \times \mathbb{R}^p$,
- $p(x)$ – полином с *положительными* коэффициентами.

Для каждой из перечисленных ниже функций K укажите, является ли она ядром. Если является, то докажите это, если нет – приведите контрпример.

(a) $K(x,z) = K_1(x,z) + K_2(x,z)$ [0.25 баллов]
(b) $K(x,z) = K_1(x,z) - K_2(x,z)$ [0.25 баллов]
(c) $K(x,z) = aK_1(x,z)$ [0.25 баллов]
(d) $K(x,z) = -aK_1(x,z)$ [0.25 баллов]
(e) $K(x,z) = K_1(x,z)K_2(x,z)$ [2 балла]
(f) $K(x,z) = f(x)f(z)$ [1 балл]
(g) $K(x,z) = K_3\left(\phi(x),\phi(z)\right)$ [1 балл]
(h) $K(x,z) = pK_1(x,z)$ [1 балл]

Подсказка: Функция в вопросе (e) в действительности является ядром. Вам необходимо это доказать. Этот результат может оказаться полезным для решения некоторых других вопросов.