1.a. Let S be a finite set of points $\{x_1, \ldots, x_\ell\}$, and let K_1' and K_2', be the corresponding kernel matrices obtained by restricting K_1 and K_2 to these points. Consider any vector $\alpha \in R^\ell$. (a matrix K is positive semi-definite if and only if $\alpha^T K \alpha \geq 0$ for all α.) We have

$$\alpha^T (K_1' + K_2') \alpha = \alpha^T K_1' \alpha + \alpha^T K_2' \alpha \geq 0 \quad (1)$$

and so $K_1' + K_2'$ is positive semi-definite and $K_1(x,z) + K_2(x,z)$ is a kernel.

1.b No, $\exists\, \alpha \in R^\ell : \alpha^T K_1' \alpha - \alpha^T K_2' \alpha < 0$.

1.c. $\alpha^T a K_1' \alpha = a \alpha^T K_1' \alpha \geq 0$,

$a K_1(x,z)$ is a kernel.

1.d. No, $\alpha^T (-a) K_1' \alpha = -a \alpha^T K_1' \alpha < 0$.

1.e. Yes, positive scalar doesn't change the vector space on which K_1 is defined, rescales inputs without inverting the sign.

$K'_{1_{ij}} = K_1(a \cdot x_i, b \cdot x_j)$, where

$a, b \in R^+$, positive real numbers.

$\forall \alpha \in R^{\ell} : \alpha^T K'_1 \alpha \geq 0.$

1.f. Yes, by closure properties of kernel functions.

1.g. Yes, by closure properties of kernel functions

1.h. Yes, by closure properties

1.i. Yes.

1.j. No

1.k. No

ЧИСЛО МІСЯЦЬ 20 РІК

ПН ВТ СР ЧТ ПТ СБ НД

2.a. Linear: $K(\varphi_2(\varphi_1(x)), \varphi_2(\varphi_1(z))) = (A_2 \cdot A_1 \cdot x)^T \cdot (A_2 \cdot A_1 \cdot z)$

2.b. Polynomial: $K(\varphi_2(\varphi_1(x)), \varphi_2(\varphi_1(z))) = ((x^T z + c_1)^{d_1} + c_2)^{d_2}$

2.c. RBF: $K(\varphi_2(\varphi_1(x)), \varphi_2(\varphi_1(z))) = \exp(2\epsilon \cdot (\exp(-\epsilon(x-z)^2) - 1))$

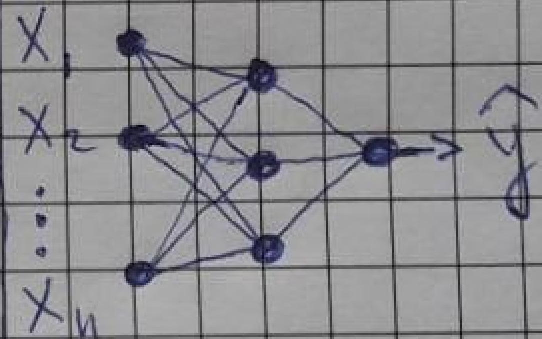
3. $$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} = \frac{e^{x} + e^{-x} - 2e^{-x}}{e^{x} + e^{-x}} =$$

$$= 1 + \frac{-2e^{-x}}{e^{x} + e^{-x}} = 1 - \frac{2}{e^{2x} + 1} = 1 - 2\sigma(-2x) =$$

$$= 1 - 2(1 - \sigma(2x))$$, by symmetry of σ around the origin

$$= 2\sigma(2x) - 1; \quad (*)$$

Let the first network computes:



$$f_1(x) = W_1^{[2]} \cdot \sigma(W_1^{[1]} x + b_1^{[1]}) + b_1^{[2]}$$

Second network computes:

$$f_2(x) = W_2^{[2]} \cdot \tanh(W_2^{[1]} x + b_2^{[1]}) + b_2^{[2]}$$

Using (*):

$$f_2(x) = 2 \cdot W_2^{[2]} \cdot \sigma(2(W_2^{[1]} \cdot x + b_2^{[1]})) - W_2^{[2]} + b_2^{[2]}$$

$f_1(x) = f_2(x)$, when
$$\begin{cases} W_1^{[1]} = 2W_2^{[1]} \\ b_1^{[1]} = 2b_2^{[1]} \\ W_1^{[2]} = 2 \cdot W_2^{[2]} \\ b_1^{[2]} = -W_2^{[2]} + b_2^{[2]} \end{cases}$$