

Exer 1b : to show convolution fce is positive $\textcircled{1}$
and semi-definite.

we need to prove : $\sum_i \sum_j d_i d_j K(x_i, x_j) \geq 0$

$$\begin{aligned} & \sum_i \sum_j d_i d_j K(x_i, x_j) \\ &= \sum_i \sum_j d_i d_j \|x_i - x_j\|^2 \end{aligned}$$

we set $A = (d_1, d_2, \dots, d_L)^T$

$$C = \begin{pmatrix} \|x_1 - x_1\|^2 & \dots & \|x_1 - x_L\|^2 \\ \|x_2 - x_1\|^2 & \dots & \|x_2 - x_L\|^2 \\ \vdots & \ddots & \vdots \\ \|x_L - x_1\|^2 & \dots & \|x_L - x_L\|^2 \end{pmatrix} = \begin{pmatrix} C^{\frac{1}{2}} \\ \vdots \\ C^{\frac{1}{2}} \end{pmatrix}^T C^{\frac{1}{2}} = \begin{pmatrix} C^{\frac{1}{2}} \\ \vdots \\ C^{\frac{1}{2}} \end{pmatrix}^T C^{\frac{1}{2}}$$

because it is a

we set ~~the~~ the general ~~inverse~~ $L \times L$ matrix ~~matrix~~ of C

$$= A^T C A$$

$$= A^T (C^{\frac{1}{2}})^T C^{\frac{1}{2}} A$$

$$= \|C^{\frac{1}{2}} A\|^2$$

the feature map can be $K(x, x') = \|x - x'\|^2$

Exer 2.a).

$$\begin{aligned} & \sum_{i=1}^K \sum_{j=1}^K d_i d_j K(x_i, x_j) \\ &= \sum_{i=1}^K \sum_{j=1}^K d_i d_j \sum_{m=1}^M \beta_m \sum_{n=1}^{N-m+1} \mathbb{1}(U_{m,n}(x_i) = U_{m,n}(x_j)) \\ &= \sum_{m=1}^M \beta_m \sum_{n=1}^{N-m+1} \sum_{i=1}^K \sum_{j=1}^K d_i d_j \mathbb{1}(U_{m,n}(x_i) = U_{m,n}(x_j)) \\ & \quad \text{or} \\ & \quad \geq 0 \end{aligned}$$

under certain m, n .

$$\sum_{i=1}^K \sum_{j=1}^K d_i d_j \mathbb{1}(U_{m,n}(x_i) = U_{m,n}(x_j))$$

we set $A = (d_1, \dots, d_K)^T$

and $C = \begin{pmatrix} & \\ & \end{pmatrix}$
 $K \times K$ matrix

$$C_{ij} = 1 \text{ if } U_{m,n}(x_i) = U_{m,n}(x_j) \\ = 0 \text{ otherwise.}$$

because C is $K \times K$ matrix
 we can find $C^{\frac{1}{2}}$ so that

$$C = (C^{\frac{1}{2}})^T C^{\frac{1}{2}}$$

$$= A^T C A$$

$$= A^T (C^{\frac{1}{2}})^T C^{\frac{1}{2}} A$$

$$= \|C^{\frac{1}{2}} A\|^2 \geq 0.$$

positive

so $\sum_{i=1}^K \sum_{j=1}^K d_i d_j K(x_i, x_j) \geq 0$. it is semi definite.

b) $M=1$.

$$K(x, x') = \beta_1 \sum_{n=2}^N \mathbb{1}(U_{1,n}(x) = U_{1,n}(x'))$$

it is the total amount of "same single gene code on the same position in two gene sequences", weighted by β_1 are the same.

c) $M=2$, $\beta_1=0$, $\beta_2=1$

$$K(x, x') = \sum_{n=2}^{N-1} \mathbb{1}(U_{2,n}(x) = U_{2,n}(x'))$$

it is the total amount of "two continuous gene code are the same on the same position on two gene sequences"