Necessary condition for matrix-valued kernels

Miaoyan Wang, April 26, 2020

Theorem 0.1 (Necessary condition)**.** *Suppose $\boldsymbol{K}\colon \mathbb{R}^{d'\times d}\times\mathbb{R}^{d'\times d}\mapsto\mathbb{R}^{d\times d}$ is a function that takes as input a pair of matrices and produces a matrix. Let $\{\boldsymbol{X}_i\in\mathbb{R}^{d'\times d}\colon i\in[n]\}$ denote a set of input matrices, and let $\mathcal{K}$ denote an order-4 (n,n,d,d)-dimensional array,*

$$\mathcal{K}=[\![\mathcal{K}(i,i',p,p')]\!],\quad \textit{where } \mathcal{K}(i,i',p,p') \textit{ is the } (p,p')\textit{-th entry of the matrix } \boldsymbol{K}(\boldsymbol{X}_i,\boldsymbol{X}_{i'}).$$

Then, the factorization $\boldsymbol{K}(\boldsymbol{X}_i,\boldsymbol{X}_{i'})=\boldsymbol{h}(\boldsymbol{X}_i)^T\boldsymbol{h}(\boldsymbol{X}_{i'})$ exists for some mapping $\boldsymbol{h}$, only if both of the following conditions hold:

(1) For every index $i\in[n]$, the matrix $\mathcal{K}(i,i,\cdot,\cdot)\in\mathbb{R}^{d\times d}$ is positive semidefinite.

(2) For every index $p\in[d]$, the matrix $\mathcal{K}(\cdot,\cdot,p,p)\in\mathbb{R}^{n\times n}$ is positive semidefinite.

Proof. (1) Let $i\in[n]$ be a fixed index. For any vector $\boldsymbol{a}\in\mathbb{R}^d$,

$$\boldsymbol{a}^T\mathcal{K}(i,i,\cdot,\cdot)\boldsymbol{a}=\boldsymbol{a}^T\boldsymbol{h}(\boldsymbol{X}_i)^T\boldsymbol{h}(\boldsymbol{X}_i)\boldsymbol{a}=\langle\boldsymbol{h}(\boldsymbol{X}_i)\boldsymbol{a},\ \boldsymbol{h}(\boldsymbol{X}_i)\boldsymbol{a}\rangle=\|\boldsymbol{h}(\boldsymbol{X}_i)\boldsymbol{a}\|_2\geq 0$$

(2) Let $p\in[d]$ be a fixed index. For any vector $\boldsymbol{b}=(b_1,\ldots,b_n)^T\in\mathbb{R}^n$,

$$\begin{aligned}
\boldsymbol{b}^T\mathcal{K}(\cdot,\cdot,p,p)\boldsymbol{b}&=\sum_{ij}b_ib_j\left[\boldsymbol{h}(\boldsymbol{X}_i)^T\boldsymbol{h}(\boldsymbol{X}_j)\right]_{(p,p)}\\
&=\sum_{ij}b_ib_j\sum_k\left[\boldsymbol{h}(\boldsymbol{X}_i)\right]_{(k,p)}\left[\boldsymbol{h}(\boldsymbol{X}_j)\right]_{(k,p)}\\
&=\sum_k\left(\sum_i\left[\boldsymbol{h}(\boldsymbol{X}_i)\right]_{(k,p)}b_i\right)\left(\sum_j\left[\boldsymbol{h}(\boldsymbol{X}_j)\right]_{(k,p)}b_j\right)\\
&=\sum_k\left(\sum_i\left[\boldsymbol{h}(\boldsymbol{X}_i)\right]_{(k,p)}b_i\right)^2\geq 0
\end{aligned}$$

$\square$