# 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 tensor,*

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

*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 $\mathcal{K}(i,i,:,:)\in\mathbb{R}^{d\times d}$ is positive semidefinite.*

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

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

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

(2) Let $p\in[d]$ be a fixed index. We use $[\cdot]_{(k,p)}$ to denote the $(k,p)$-th entry of the matrix. For any vector $\boldsymbol{b}=(b_1,\ldots,b_n)^T\in\mathbb{R}^n$,

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

□

Updated on April 29, 2020. Generalization of Mercer's theorem to matrix-valued predictors.

**Theorem 0.2** (Necessary and Sufficient condition (sample version)). *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 $\mathcal{X}=\{\boldsymbol{X}_i\in\mathbb{R}^{d'\times d}\colon i\in[n]\}$ denote a set of input matrices, and let $\mathcal{K}$ denote an order-4 $(n,n,d,d)$-dimensional tensor,*

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

*Then, the following two statements are equivalent:*

- *There exists a feature mapping* $\boldsymbol{h}(\cdot)$ *over the input set* $\mathcal{X}$ *such that* $\boldsymbol{K}(\boldsymbol{X}_i, \boldsymbol{X}_{i'}) = \boldsymbol{h}(\boldsymbol{X}_i)^T \boldsymbol{h}(\boldsymbol{X}_{i'})$.
- *The matrix* $\mathcal{K}_{(13)(24)} \in \mathbb{R}^{nd \times nd}$ *is positive semidefinite. Here* $\mathcal{K}_{(13)(24)}$ *denotes the square unfolding of the tensor* $\mathcal{K}$.

**Remark 1.** Recall that earlier we have defined two types of kernel $\boldsymbol{K}$:

- Hadamard-product type: $\boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X}') = \underbrace{(\boldsymbol{X}^T \boldsymbol{X}' + \mathbb{1}\mathbb{1}^T) \circ \cdots \circ (\boldsymbol{X}^T \boldsymbol{X}' + \mathbb{1}\mathbb{1}^T)}_{d \text{ times}}$.
- Matrix-polynomial type: $\boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X}') = (\boldsymbol{X}^T \boldsymbol{X}' + \mathbb{1}\mathbb{1}^T)^d$.

Theorem 0.2 implies that the existence of feature mapping can be verified using the positive-definiteness of $\mathcal{K}_{(13)(24)}$. Straightforward calculation shows that the Hadamard-product type kernel $\boldsymbol{K}$ is valid whereas the matrix-polynomial type kernel is not.

*Proof of Theorem 0.2.* "$\Leftarrow$" Suppose $\mathcal{K}_{(13)(24)} \in \mathbb{R}^{nd \times nd}$ is a positive semidefinite matrix. Then there exists a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times (nd)}$ such that

$$\mathcal{K}_{(13)(24)} = \boldsymbol{A}^T \boldsymbol{A}. \tag{1}$$

We reshape $\boldsymbol{A}$ into an $(m, n, d)$-dimensional tensor, and with a little abuse of notation, we still use $\boldsymbol{A}$ to denote the resulting object. We define the feature mapping $\boldsymbol{h}\colon \mathbb{R}^{d' \times d} \mapsto \mathbb{R}^{m \times d}$ as follows:

$$\begin{aligned} \boldsymbol{h}\colon \mathbb{R}^{d' \times d} &\mapsto \mathbb{R}^{m \times d} \\ \boldsymbol{X}_i &\mapsto \boldsymbol{h}(\boldsymbol{X}_i) \stackrel{\text{def}}{=} \boldsymbol{A}(:, i, :), \quad \text{for all } i \in [n]. \end{aligned}$$

The defined mapping satisfies that, for all $i, i' \in [n]$,

$$\begin{aligned} \boldsymbol{h}(\boldsymbol{X}_i)^T \boldsymbol{h}(\boldsymbol{X}_{i'}) &= \boldsymbol{A}(:, i, :)^T \boldsymbol{A}(:, i', :) \\ &= \mathcal{K}(i, :, i', :) \\ &= \boldsymbol{K}(\boldsymbol{X}_i, \boldsymbol{X}_{i'}), \end{aligned}$$

where the second line follows form (1) and the third line follows from the definition of $\mathcal{K}$.

"$\Rightarrow$" Let $\boldsymbol{a} = (a_{11}, \ldots, a_{1d}, a_{21}, \ldots, a_{2d}, \ldots, a_{nd})^T$ denote an arbitrary vector in $\mathbb{R}^{nd}$. Note that

$$\begin{aligned} \boldsymbol{a}^T \mathcal{K}_{(13)(24)} \boldsymbol{a} &= \sum_{i, i', p, p'} a_{ip} \mathcal{K}(i, i', p, p') a_{i'p'} \\ &= \sum_{i, i', p, p'} a_{ip} \left[\boldsymbol{K}(\boldsymbol{X}_i, \boldsymbol{X}_{i'})\right]_{pp'} a_{i'p'} \\ &= \sum_{i, i', p, p'} a_{ip} \left( \sum_k \left[\boldsymbol{h}(\boldsymbol{X}_i)^T\right]_{pk} \left[\boldsymbol{h}(\boldsymbol{X}_{i'})\right]_{kp'} \right) a_{i'p'} \end{aligned}$$

$$= \sum_k \left( \sum_{i,p} a_{ip} \left[ \boldsymbol{h}(\boldsymbol{X}_i) \right]_{kp} \right) \left( \sum_{i',p'} a_{i'p'} \left[ \boldsymbol{h}(\boldsymbol{X}_{i'}) \right]_{kp'} \right)$$

$$= \sum_k \left( \sum_{i,p} a_{ip} \left[ \boldsymbol{h}(\boldsymbol{X}_i) \right]_{kp} \right)^2 \geq 0.$$

Therefore, $\mathcal{K}_{(13)(24)}$ is positive semi-definite. $\square$