$\mathcal{GP}$ Write Up 5

September 26, 2018

Convolutions on the Kernel matrix. Co-variance function is stationary.

1. Convolutions on a Kernel matrix naive way

Lets say we have an image $y \in \mathbb{R}^{n \times n}$. Then the kernel matrix will be of size $K \in \mathbb{R}^{n^2 \times n^2}$. We are trying to approximate for v that is $y = Kv$, where $v \in \mathbb{R}^{n \times n}$.

Lets take for example an image y with $n = 2$. Then:

$$y \approx \begin{bmatrix} k(0,0) & k(0,-1) & k(-1,0) & k(-1,-1) \\ k(0,1) & k(0,0) & k(-1,1) & k(-1,0) \\ k(1,0) & k(1,-1) & k(0,0) & k(0,-1) \\ k(1,1) & k(1,0) & k(0,1) & k(0,0) \end{bmatrix} \begin{bmatrix} v_{0,0} \\ v_{0,1} \\ v_{1,0} \\ v_{1,1} \end{bmatrix} \tag{1}$$

To implement convolutions in the naive way we can consider a filter of the size $\mathbb{R}^{1 \times 4}$, with stride as one and zero padding, the convolution goes over a row in one stride. The computational complexity of this method is $O(n^4)$.

2. Convolutions on a Kernel

If we look at the above kernel we see that some elements $k(i, j)$ are repeated. We can avoid these and write the kernel matrix like so.

$$K' = \begin{bmatrix} k(-1,-1) & k(-1,0) & k(-1,1) \\ k(0,-1) & k(0,0) & k(0,1) \\ k(1,-1) & k(1,0) & k(1,1) \end{bmatrix} \tag{2}$$

and

$$v = \begin{bmatrix} v_{1,1} & v_{1,0} \\ v_{0,1} & v_{0,0} \end{bmatrix} \tag{3}$$

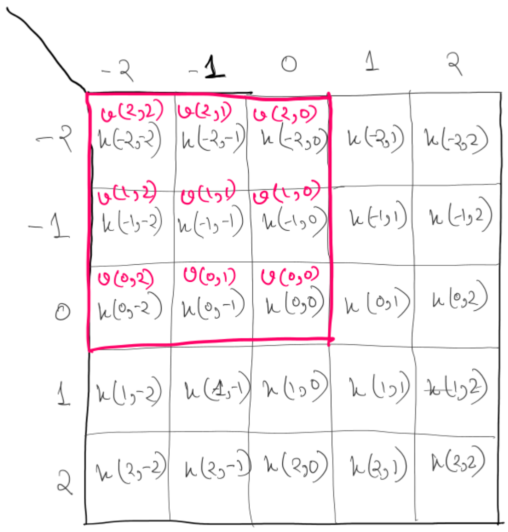
and we say

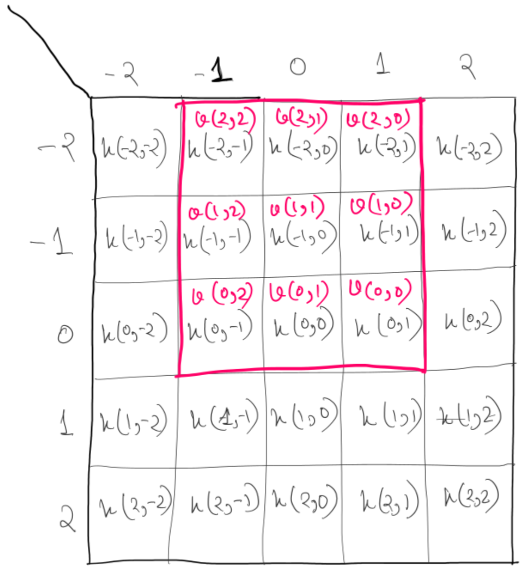
$$y \approx K' * v \tag{4}$$

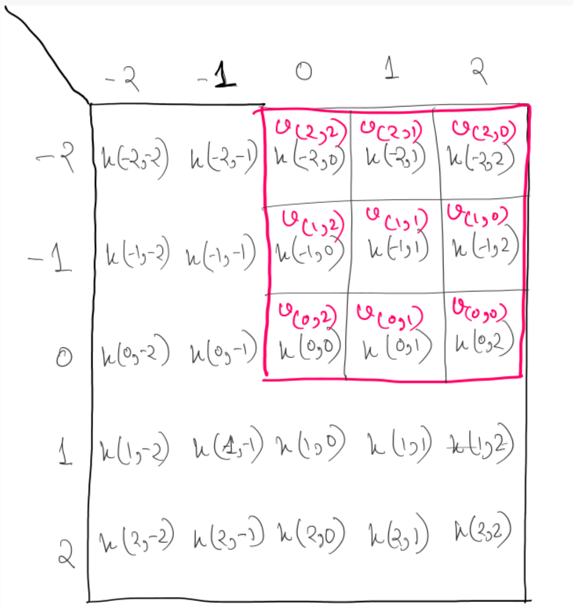
For any $y \in \mathbb{R}^{n \times n}$ we have $K' \in \mathbb{R}^{(2n-1) \times (2n-1)}$.

The following is an illustration for $y \in \mathbb{R}^{3 \times 3}$, $K' \in \mathbb{R}^{5 \times 5}$ and $v \in \mathbb{R}^{3 \times 3}$. This does not need any zero padding and we have to perform the above convolution with stride one.

$$v = \begin{bmatrix} v_{2,2} & v_{2,1} & v_{2,0} \\ v_{1,2} & v_{1,1} & v_{1,0} \\ v_{0,2} & v_{0,1} & v_{0,0} \end{bmatrix} \tag{5}$$







The computational complexity is still $O(n^4)$ however we have decreased the space complexity from $O(n^4)$ to $O(n^2)$.

From the above illustrations we can see that we **don't have to pad** the kernel with zeros. However if we make a stronger assumption of the co-variance function being *isotropic*, do we have to pad with **zeros**?

With isotropic covariance function we can simplify the kernel matrix as:

$$K_{iso} = \begin{bmatrix} k(0,0) & k(0,1) & k(0,2) \\ k(1,0) & k(1,1) & k(1,2) \\ k(2,0) & k(2,1) & k(2,2) \end{bmatrix} \tag{6}$$