Derivation of the effect of symmetry

Suppose we consider the covariance between the k entries of $\mathbf{y}$ in a Gaussian process. Assume the covariance matrix is generated by some covariance function $C(\mathbf{x}, \mathbf{x}')$ that depends only on the distance between $\mathbf{x}$ and $\mathbf{x}'$, i.e. $C(\mathbf{x}, \mathbf{x}') = C(|\mathbf{x} - \mathbf{x}'|)$. Then the covariance matrix $\mathbf{C}$ is a Toeplitz matrix. Suppose we have a set of k points $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ and we wish to find the covariance between the sum of the k points and another point $\mathbf{x}_0$. Then we have

$$\mathrm{cov}(y(\mathbf{x}_0), \sum_{i=1}^{k} y(\mathbf{x}_i)) = \sum_{i=1}^{k} \mathrm{cov}(y(\mathbf{x}_0), y(\mathbf{x}_i)) = \sum_{i=1}^{k} C(|\mathbf{x}_0 - \mathbf{x}_i|)$$

The same sum is obtained for a permutation of the ordering of the k points. Thus, if we have a set of n points $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ and we ask for the covariance between the sum of k points and another point $\mathbf{x}_0$, then we have

$$\mathrm{cov}(y(\mathbf{x}_0), \sum_{i=1}^{k} y(\mathbf{x}_i)) = \frac{k!(n-k)!}{n!} \sum_{\text{all } k \text{ subsets}} \sum_{i=1}^{k} C(|\mathbf{x}_0 - \mathbf{x}_i|)$$

where the second sum is taken over all possible k-subsets of the n points. Now, consider the sum

$$\sum_{\text{all } k \text{ subsets}} \sum_{i=1}^{k} C(|\mathbf{x}_0 - \mathbf{x}_i|)$$

and the sum

$$\sum_{\text{all } k \text{ subsets}} \sum_{i=1}^{k} C(|\mathbf{x}_0 - \mathbf{x}_{n-i}|)$$

where $\mathbf{x}_{n-i}$ is the $(n-i)$th point in the ordering of the n points. Then, if the distribution of points is symmetric, these two sums are equal. Thus, if we have a symmetric distribution of points, then

$$\mathrm{cov}(y(\mathbf{x}_0), \sum_{i=1}^{k} y(\mathbf{x}_i)) = \mathrm{cov}(y(\mathbf{x}_0), \sum_{i=1}^{k} y(\mathbf{x}_{n-i}))$$

Thus, the covariance between $y(\mathbf{x}_0)$ and the sum of the k points is equal to the covariance between $y(\mathbf{x}_0)$ and the sum of the k points furthest away from $\mathbf{x}_0$. Now, consider the sum

$$\sum_{\text{all } k \text{ subsets}} \sum_{i=1}^{k} \left[C(|\mathbf{x}_0 - \mathbf{x}_i|) + C(|\mathbf{x}_0 - \mathbf{x}_{n-i}|) \right]$$

Then, if the distribution of points is symmetric, we have

$$\sum_{\text{all } k \text{ subsets}} \sum_{i=1}^{k} \left[C(|\mathbf{x}_0 - \mathbf{x}_i|) + C(|\mathbf{x}_0 - \mathbf{x}_{n-i}|) \right] = 2 \sum_{\text{all } k \text{ subsets}} \sum_{i=1}^{k} C(|\mathbf{x}_0 - \mathbf{x}_i|)$$

Thus, if the distribution of points is symmetric, then the covariance between $y(\mathbf{x}_0)$ and the sum of the k points is equal to one half the sum of the covariances between $y(\mathbf{x}_0)$ and the k points and the k points furthest away from $\mathbf{x}_0$. Thus, if the distribution of points is symmetric, then the covariance between $y(\mathbf{x}_0)$ and the sum of the k points is equal to the average covariance between $y(\mathbf{x}_0)$ and k points.