Simple Overview of Statistical Concepts of week 8

Monitoring network optimization. In the context of monitoring network optimization, spatial kriging can be used to identify the locations where stations/observers should be placed or removed to obtain the most accurate representation/prediction of the data in the area. The technique involves selecting a set of candidate locations and then using spatial kriging to estimate the values at those locations based on the data collected at the existing locations. The locations that contribute to the highest estimation errors are then prioritized for the placement/removal of new monitoring stations.

Monitoring network optimization using spatial kriging with minimum mean-variance criteria minimizes the mean-variance of the estimation error in predicting spatial variability in a given area. The stations that have the smallest mean-variance increase are first candidates for removal as they will not substantially impact the quality of spatial predictions.

Another important practical aim is to estimate values at unobserved locations and use them to create contour maps that visualize the spatial distribution. Contrary to the mean-variance approach, monitoring network optimization for delineating contours removes stations that have the largest deviations from the expected contour lines as they will not substantially impact the placement of contour lines.

Spatal point processes. Spatial point processes model the occurrence of random points in a spatial region. In spatial point processes, the intensity function plays a crucial role. The intensity function represents the expected number of points per unit area in a given spatial region. The intensity function is often denoted by λ and is a function of the spatial coordinates, for example, (x, y).

There are different types of intensity functions that can be used to model spatial point processes, depending on the characteristics of the process. Two common types of intensity functions include:

- Homogeneous intensity function: This is a constant intensity function that assumes that the process is stationary and have same statistical properties at different locations of the spatial region.
- Inhomogeneous intensity function: This is a non-constant intensity function that allows for spatial changes in the distribution of points.

Kernel estimator. The kernel estimator is a method for estimating the spatial intensity function of a point process. The kernel function is a smooth function that decays with distance from its center. It determine weights/impact of the existing observations for the estimated intensity at a particular location. The closer a point is to the reference point, the greater its weight. The kernel estimator places a kernel function at each observed point in the point process. Then, the estimated spatial intensity function is obtained by summing the kernel functions at each location in space and dividing by the total number of points in the point process. The resulting estimate is a smooth function that describes the expected number of events per unit area in different parts of the region.

There are different choices of kernel functions that can be used in the kernel estimator. A commonly used kernel function is the Gaussian kernel. It uses a parameter sigma to control the impact of points depending on distances. When sigma is larger, the Gaussian kernel assigns significant weight to points that are further away from the reference point. This can result in over-smoothing of the data and a reduction in the ability of the kernel to capture local variations in the data. On the other hand, when sigma is smaller, the Gaussian kernel assigns more weight to nearby points, resulting in a sharper kernel that may picks out only existing points and ignore the overall trend. The choice of sigma can be determined visually or using cross-validation or other optimization techniques to find the value that yields the best performance on a particular task.

Spatial covariates.

Spatial covariates refer to the explanatory variables that are spatially distributed and may influence the response variable. Spatial covariates may include factors such as land use, climate, topography, and socio-economic characteristics of an area. The inclusion of spatial covariates in spatial models can improve the accuracy of the estimates and provide useful practical interpretations of spatial data patterns.