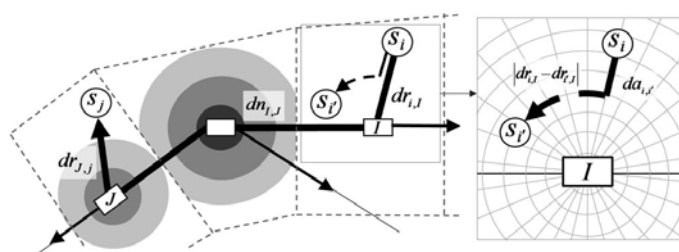


## Development of kriging by introducing covariance function of shortest-path distance

Daisuke Murakami<sup>1\*</sup>, Morito Tsutsumi<sup>1</sup>

<sup>1</sup>University of Tsukuba

Generally, spatial data obtained from neighboring locations correlate well, whereas those from widely separated sites are less so. This property is called spatial dependence and is an important feature when handling spatial data, e.g., when performing spatial interpolation. In geostatistics, Kriging is a key technique used for spatial interpolation. In



Definition of distance measures and image of spatial dependence

traditional geostatistics, covariance is modeled by using a function of Euclidean distance, called the covariance function, to express spatial dependence. However, Euclidean distance may not fully describe spatial dependence in real cities or regions linked by transportation networks; the linking strongly affects spatial dependence. In such case, the enforcement of spatial dependence may exist through the nodes of transportation networks and the spill-over spread not only along the networks/links but also across nodes throughout the city/region.

In order to consider the effects of networks such as railways or air networks explicitly, it is quite natural to introduce the shortest-path distance. Thus, in this study, we propose a new covariance function considering the shortest-path distance and develop a spatial interpolation technique that is theoretically consistent with conventional kriging. The networks that we consider are transportation networks where people enter and exit at nodes, such as highways, railways, and air networks. To explicitly consider the spatial dependence resulting from the existence of networks, we divide the entire space into domains by employing spatial tessellation based on nodes. We assume that spatial dependence is enforced by the socioeconomic activities that are affected by the traffic flowing through the nodes. In this study, we adopt Voronoi tessellation (diagram). First, the study presents a technique to explicitly consider the spatial dependency resulting from the existence of transportation networks. To be more precise, spatial tessellation based on nodes is applied. It is assumed that spatial dependence is enforced by the socioeconomic activities that are affected by the traffic flow through the nodes and the spill-over effect in each spatial domain linked by the nodes.

Next, the study introduces a non-Euclidean distance, more precisely, the shortest-path distance in a network satisfying the distance axiom. In order to formulate a new covariance function, we define three distance measures, as shown in the figure: the distance between nodes along the network ( $dn$ ); distance in radial direction ( $dr$ ), that is, gap between the distances to the nearest node in the domain; and the arc distance on a circle centered at a node in the domain ( $da$ ).

Subsequently, the study presents the exponential covariance function of the non-Euclidean distance that ensures positive definiteness, by using the concept of extended shortest path tree. However, it is natural to consider that the Euclidean distance also affects spatial dependence as conventional kriging does. Therefore, the study demonstrates a method to combine both Euclidean - and non-Euclidean-distance-based covariance functions.

Finally, the study examines the applicability of the proposed technique. For this, we apply it to a case study of land price interpolation in Tokyo Metropolitan Area and compare the results with those obtained by conventional kriging. We use the officially assessed land price data for the year 2007 provided by the Ministry of Land, Infrastructure and Transport, Japan. The case study area is Tokyo Metropolitan Area. Empirical results show that the proposed technique is better than conventional kriging, both in terms of model fitting and prediction accuracy. Furthermore, parameter estimation by the proposed technique shows that half of the spatial dependence could be explained satisfactorily by shortest-path distances rather than Euclidean distances.

Keywords: spatial interpolation, kriging, shortest path distance, network, covariance function, positive definite