Abstract

We aim to carry out effective clustering of geographically distributed time series data. This work uses singular value decomposition (SVD) for dimension reduction and spectral clustering. This model approach provides insight that there may be a common structure to disease spread, which contradicts early common narratives about how SARS-CoV-2 spread. In particular, our model indicates greater normalized rates of spread in rural areas, rather than the urban centers and transportation hubs that were widely thought to have higher rates of infection during the pandemic. Our novel method provides a temporal evolution of clusters derived from sliding window SVDs via spectral clustering. When applied to SARS-CoV-2 case load data, we determine geo-temporal patterns of disease spread. The method can be a tool to help an expert discern and interpret various patterns of disease spread. Wepresent applications to the disease datasets from the US counties and Italian regions.