

[Tuan Nguyen kcstevone@gmail.com](mailto:kcstevone@gmail.com)

**ABSTRACT:** Dimension reduction methods have come to the forefront of many applications where the number of covariates,  $p$ , far exceeds the sample size,  $N$ . For example, in survival analysis studies using microarray gene expression data, 10-30K expressions per patient are collected, but only a few hundred patients are available for the study. Due to the high dimensionality of the microarray dataset, one needs to reduce the dimension of the gene expression data matrix before embarking on any type of analysis. One approach to do this is using dimension reduction methods. In this talk, I will focus on several well-known linear dimension reduction methods in the literature. In particular, I will discuss two methods explored in my thesis work: the method of Random Projection and the method of Rank-based Partial Least Squares.

In Random Projection (RP), the original  $p$ -dimensional data matrix is projected onto a  $k$ -dimensional subspace using a random matrix. The main motivation of RP is the Johnson-Lindenstrauss (JL) Lemma which uses the criterion of preserving pairwise  $L_2$  distances among the points for projecting points from high to low-dimensional space within 1 plus or minus a small distortion. In this talk, I will discuss our improvements for the lower bound for  $k$  when using pairwise  $L_2$  ( $L_2$ ) distances in the space of the original points and pairwise  $L_2$  ( $L_1$ ) distances in the space of the projected points, respectively, in the cases when the random matrix is standard Gaussian and of Achlioptas-type. I will also discuss our proposed dimension reduction method of Rank-based Modified Partial Least Squares (RMPLS), which works well in the presence of outliers. The weight vectors of RMPLS can be derived as a solution of an optimization problem. Finally, I will present some comparisons of the different dimension reduction methods using survival analysis with microarray data based on simulations and real datasets.