

- **Project Title:** The Cox Regression of Microarray Gene Expression Data Using Correlated Principal Component Analysis

**Investigator:** Qiang Zhao

**Department:** Mathematics

**Project Summary:**

In the past decade, microarray technology has been developed and become widely used in biology, medical sciences, and agriculture etc. With this new technology, scientists can study thousands of genes simultaneously by measuring their expression levels and determine their functionality or effect on a certain phenotype, such as disease development or survival time of a patient. However, the analysis of the resulting gene expression data faces two challenges. First, high dimensionality of gene expression data makes invalid of the conventional regression analysis. Second, the survival time is subject to censoring so that the survival time of a patient may be only known partially.

In this project, in order to predict the survival of a patient based on the patients gene expression information, we proposed a new approach, which is a combination of the Cox survival analysis and the correlation principal component analysis. The proposed approach first selects components of the gene expression data that are related to the survival and then uses them in the Cox regression model to handle the censoring problem of survival data. The results based on simulated data and a set of publicly available data on diffuse large B-cell lymphoma show that the proposed method works well in terms of models' robustness and predictive ability in comparison with some existing partial least squares approaches. Also, the new approach is simpler and easy to implement.

**Publications:**

"Cox Survival Analysis of Microarray Gene Expression Data Using Correlation Principal Component Regression", *Statistical Applications in Genetics and Molecular Biology*: Vol. 6: Iss. 1, Article 16 (2007)

**Presentations:**

"Survival Analysis of Microarray Gene Expression Data Using Correlation Principal Component Regression", The third Erich L. Lehmann Symposium, Houston, Texas, May 16-19, 2007.