



## 성균관대학교 응용통계연구소 세미나 안내

응용통계연구소 2월 정기 세미나를 다음과 같이 개최합니다.

- 일 시: 2012년 2월 9일 목요일 오전 10:00 - 11:30
- 장 소: 경제관 2층 32211호

Dr. Chun-houh Chen

- 발표제목: Exploratory Data Analysis for Symbolic Data with Matrix Visualization

Generalized Association Plots (GAP) for matrix visualization (MV) has obtained encouraging results for continuous and categorical data for assisting conventional methods in producing more comprehensive analysis works. This project intends to extend matrix visualization technology to related research fields of Symbolic Data Analysis (SDA). Many graphical and visualization tools have been developed by related researchers for SDA. Somehow all these tools also inherited most of the limitations and disadvantages from their counter-parts for conventional statistical analysis. With appropriate selections (development) of proximity matrices for symbolic concepts/variables and suitable displaying mechanism for symbolic units, we can borrow techniques from original GAP environment to permute (or cluster) raw data matrix with two proximity matrices and to display matrix maps for symbolic data. Through suitable displaying techniques with seriations, it is hoped the clusters of symbolic units, the grouping structure of variables, and interactions embedded in the symbolic data set can be visually extracted. It will be very useful for exploration and understanding of symbolic data structure even with huge dataset.

## Dr. Junji Nakano

- 발표제목: Aggregated symbolic data analysis

Symbolic Data Analysis (SDA) handles symbolic data (SD), in which values of a variable can be more complex than the traditional data such as real numbers and categorical values. Typical SD take intervals, histograms and bar charts as variable values (Billard and Diday, 2006). SDA provides techniques for handling such SD, including several extensions of principle component analysis (PCA) and regression analysis.

In this talk we propose to use correlation coefficients among variables in each SD for several SDA. We notice that SD often arise by aggregation of individuals in groups. In this situation, correlation coefficients are easily calculated together with traditional SD information which describes information about the marginal distribution of each variable, and are naturally used in some analyses of SD such as data visualization, PCA and regression analysis.

## 성균관대학교 응용통계연구소

Homepage: <http://stat.skku.ac.kr/rias/>

☎ 02-760-1295

e-mail: [riasmaster@skku.edu](mailto:riasmaster@skku.edu)