

# Doubly constrained factor models with applications

Henghsiu Tsai<sup>1</sup> (蔡恆修), Ruey S. Tsay<sup>2</sup> (蔡瑞胸),  
Edward M. H. Lin<sup>1\*</sup> (林孟樺) and Ching-Wei Cheng<sup>3</sup> (鄭經維)

<sup>1</sup>Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

<sup>2</sup>Booth School of Business, University of Chicago, Illinois, USA

<sup>3</sup>Department of Statistics, Purdue University, Indiana, USA

## Abstract

Factor models are widely used in the statistical literature for handling high-dimensional problems. In most applications, factors are latent variables and must be inferred from the observed data. Identifying common factors, however, is not straightforward when the dimension is high. On the other hand, an analyst may have some prior knowledge, vague or precise, about the factors. Incorporating such prior information can often lead to simplification in statistical inference. In this paper, we consider a doubly constrained factor model that enables analysts to specify both row and column constraints of the data matrix to sharpen the estimation of common factors. The row constraints may represent classifications of individual subjects whereas the column constraints may show the grouping of variables. We derive both the maximum likelihood and least squares estimates of the proposed doubly constrained factor model and use simulation to study the performance of the analysis in finite samples. Akaike information criterion is used to perform model selection. Monthly U.S. housing starts of nine different divisions are used to demonstrate the application of the proposed model.

Keywords: Akaike information criterion, factor model, constrained factor model, principal component analysis, eigenvalues, non-nested hypothesis