Mach Learn (2013) 91:279–303
DOI 10.1007/s10994-013-5345-8

Learning a factor model via regularized PCA

Yi-Hao Kao · Benjamin Van Roy

Received: 15 July 2012 / Accepted: 30 March 2013 / Published online: 20 April 2013

Abstract We consider the problem of learning a linear factor model. We propose a regularized form of principal component analysis (PCA) and demonstrate through experiments with synthetic and real data the superiority of resulting estimates to those produced by pre-existing factor analysis approaches. We also establish theoretical results that explain how our algorithm corrects the biases induced by conventional approaches. An important feature of our algorithm is that its computational requirements are similar to those of PCA, which enjoys wide use in large part due to its efficiency.

Keywords Principal component analysis · Factor model · High-dimensional data · Covariance matrix estimation · Regularization

1 Introduction

Linear factor models have been widely used for a long time and with notable success in economics, finance, medicine, psychology, and various other natural and social sciences (Harman 1976). In such a model, each observed variable is a linear combination of unobserved common factors plus idiosyncratic noise, and the collection of random variables is jointly Gaussian. We consider in this paper the problem of learning a factor model from a training set of vector observations. In particular, our learning problem entails simultaneously estimating the loadings of each factor and the residual variance of each variable. We seek an estimate of these parameters that best explains out-of-sample data. For this purpose, we consider the likelihood of test data that is independent of the training data. As such, our goal is to design a learning algorithm that maximizes the likelihood of a test set that is not used in the learning process.

Editor: Csaba Szepesvari.

Y.-H. Kao (✉) · B. Van Roy
Stanford University, Stanford, CA, USA
e-mail: yhkao@alumni.stanford.edu

B. Van Roy
e-mail: bvr@stanford.edu