Han, Yuefeng; Chen, Rong; Zhang, Cun-Hui

**Rank determination in tensor factor model.** (English) Zbl 07524962

Electron. J. Stat. 16, No. 1, 1726-1803 (2022)

Summary: Factor model is an appealing and effective analytic tool for high-dimensional time series, with a wide range of applications in economics, finance and statistics. This paper develops two criteria for the determination of the number of factors for tensor factor models where the signal part of an observed tensor time series assumes a Tucker decomposition with the core tensor as the factor tensor. The task is to determine the dimensions of the core tensor. One of the proposed criteria is similar to information based criteria of model selection, and the other is an extension of the approaches based on the ratios of consecutive eigenvalues often used in factor analysis for panel time series. Theoretically results, including sufficient conditions and convergence rates, are established. The results include the vector factor models as special cases, with an additional convergence rates. Simulation studies provide promising finite sample performance for the two criteria.

**MSC:**

- 62H25 Factor analysis and principal components; correspondence analysis
- 62H12 Estimation in multivariate analysis
- 62F07 Statistical ranking and selection procedures

**Keywords:**

- high-dimensional tensor data; factor model; rank determination; eigenvalues; Tucker decomposition

**Software:**

Cross; FinTS

**Full Text:** DOI Link

**References:**

[1] Ahn, S. C. and Horenstein, A. R. (2013). Eigenvalue ratio test for the number of factors. \textit{Econometrica} 81 1203-1227.
- Zbl 1274.62403 · doi:10.3982/ECTA8968

[2] ALESSI, L., BARIGOZZI, M. and CAPASSO, M. (2010). Improved penalization for determining the number of factors in approximate factor models. \textit{Statistics \& Probability Letters} 80 1806-1813.
- Zbl 1202.62081

[3] AMENGUAL, D. and WATSON, M. W. (2007). Consistent estimation of the number of dynamic factors in a large \((N)\) and \((T)\) panel. \textit{Journal of Business \& Economic Statistics} 25 91-96.

[4] Bai, J. (2003). Inferential theory for factor models of large dimensions. \textit{Econometrica} 71 135-171.
- Zbl 1136.62354 · doi:10.1111/1468-0262.00392

[5] BAI, C., LI, Q. and OUYANG, M. (2014). Property taxes and home prices: A tale of two cities. \textit{Journal of Economicometrics} 180 1-15.
- Zbl 1298.91122

[6] Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. \textit{Econometrica} 70 191-221.
- Zbl 1103.91399 · doi:10.1111/1468-0262.00273

[7] BAI, J. and NG, S. (2007). Determining the number of primitive shocks in factor models. \textit{Journal of Business \& Economic Statistics} 25 52-60.

[8] Barabasi, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. \textit{Science} 286 509-512.
- Zbl 1226.05223 · doi:10.1126/science.286.5439.509

[9] BRADLEY, R. C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. \textit{Probability Surveys} 2 107-144.
- Zbl 1189.60077

[10] BROWN, S. J. (1989). The number of factors in security returns. \textit{The Journal of Finance} 44 1247-1262.

[11] CAI, T. T., HAN, X. and PAN, G. (2020). Limiting laws for divergent spiked eigenvalues and largest nonspiked eigenvalue of sample covariance matrices. \textit{The Annals of Statistics} 48 1255-1280.
- Zbl 1456.62115

[12] Chamberlain, G. and Rothschild, M. (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets. \textit{Econometrica} 51 1281-1304.
- Zbl 0523.90017 · doi:10.1353/qam.1992.1190

[13] CHEN, E. Y. and CHEN, R. (2019). Modeling dynamic transport network with matrix factor models: with an application to
CHEN, E. Y., XIA, D., CAI, C. and FAN, J. (2020). Semiparametric tensor factor analysis by iteratively projected SVD. \textit{arXiv preprint arXiv:2007.02404.}

Fan, J., Liao, Y. and Mincheva, M. (2011). High dimensional covariance matrix estimation in approximate factor models. \textit{The Annals of Statistics} 39 3320. - Zbl 1246.62015.

FAN, J., WANG, W. and ZHONG, Y. (2019). Robust covariance estimation for approximate factor models. \textit{Journal of Econometrics} 208 5-22. - Zbl 1452.62410.

FAN, J. and YAO, Q. (2003). \textit{Nonlinear time series: Nonparametric and parametric methods}. \textit{Springer Series in Statistics}. Springer-Verlag, New York.

FON, M. and REICHLIN, L. (1998). Let’s get real: a factor analytical approach to disaggregated business cycle dynamics. \textit{The Review of Economic Studies} 65 453-473. - Zbl 0911.90007.

FORNI, M., HALLIN, M., LIPPI, M. and REICHLIN, L. (2005). The generalized dynamic factor model: one-sided estimation and forecasting. \textit{Journal of the American Statistical Association} 100 830-840. - Zbl 1172.62334.

FORNI, M., GIANNONE, D., LIPPI, M. and REICHLIN, L. (2006). Does information help recovering structural shocks from past observations? \textit{The European Economic Review} 50 1367-1388.

HAFNER, C. M., LINTON, O. B. and TANG, H. (2020). Estimation of a multiplicative correlation structure in the large dimensional case. \textit{Journal of Econometrics} 217 431-470. - Zbl 1456.62774.

HALLIN, M. and LIŠKA, R. (2007). Determining the number of factors in the general dynamic factor model. \textit{Journal of the American Statistical Association} 102 603-617. - Zbl 1172.62339.

HAN, Y., CHEN, R., YANG, D. and ZHANG, C.-H. (2020). Tensor factor model estimation by iterative projection. \textit{arXiv preprint arXiv:2006.02611.}

HOFF, P. D. (2011). Separable covariance arrays via the Tucker product, with applications to multivariate relational data. \textit{Bayesian Analysis} 6 179-196. - Zbl 1330.62132.

HOFF, P. D. (2015). Multilinear tensor regression for longitudinal relational data. \textit{Annals of Applied Statistics} 9 1169-1193. - Zbl 1345.62307.

JIANG, B., LI, J. and YAO, Q. (2020). Autoregressive networks. \textit{arXiv preprint arXiv:2010.04492.}

JUNG, S., LEE, M. H. and AHN, J. (2018). On the number of principal components in high dimensions. \textit{Biometrika} 105 389-402. - Zbl 06702419.

KAPETANIOS, G. (2010). A testing procedure for determining the number of factors in approximate factor models with large datasets. \textit{Journal of Business & Economic Statistics} 28 397-409. - Zbl 1214.62068.

KIM, D. and FAN, J. (2019). Factor GARCH-Itô models for high-frequency data with application to large volatility matrix prediction. \textit{Journal of Econometrics} 208 395-417. - Zbl 1452.62774.

KONG, X.-B. (2017). On the number of common factors with high-frequency data. \textit{Biometrika} 104 397-410. - Zbl 067072214.

LAM, C., YAO, Q. and BATHIA, N. (2011). Estimation of latent factors for high-dimensional time series. \textit{Biometrika} 98 901-918. - Zbl 1228.62010.

LAM, C. and YAO, Q. (2012). Factor modeling for high-dimensional time series: inference for the number of factors. \textit{The Annals of Statistics} 40 694-726. - Zbl 1273.62014.

LI, K. T. and BELL, D. R. (2017). Estimation of average treatment effects with panel data: Asymptotic theory and implementation. \textit{Journal of Econometrics} 197 65-75. - Zbl 1443.62488.

LI, H., LI, Q. and SHI, Y. (2017). Determining the number of factors when the number of factors can increase with sample size. \textit{Journal of Econometrics} 197 76-86. - Zbl 1443.62487.

LINTON, O. B. and TANG, H. (2020). Estimation of the Kronecker Covariance Model by Quadratic Form. \textit{arXiv preprint arXiv:1906.08908.}

LIU, T., YUAN, M. and ZHAO, H. (2017). Characterizing spatiotemporal transcriptome of human brain via low rank tensor decomposition. \textit{arXiv preprint arXiv:1702.07449.}

LUO, W. and LI, B. (2016). Combining eigenvalues and variation of eigenvectors for order determination. \textit{Biometrika} 105 389-402.
[45] MERLEVEDE, F., PELIGRAD, M. and RIO, E. (2011). A Bernstein type inequality and moderate deviations for weakly dependent sequences. \textit{Probability Theory and Related Fields} 151 435-474. · Zbl 1242.60020

[46] ONATSKI, A. (2010). Determining the number of factors from empirical distribution of eigenvalues. \textit{The Review of Economics and Statistics} 92 1004-1016.

[47] ONATSKI, A. (2012). Asymptotics of the principal components estimator of large factor models with weakly influential factors. \textit{Journal of Econometrics} 168 244-258. · Zbl 1443.62497

[48] OUYANG, M. and PENG, Y. (2015). The treatment-effect estimation: A case study of the 2008 economic stimulus package of China. \textit{Journal of Econometrics} 188 545-557. · Zbl 1337.62381

[49] PAN, J. and YAO, Q. (2008). Modelling multiple time series via common factors. \textit{Biometrika} 95 365-379. · Zbl 1437.62574

[50] ROSENBLATT, M. (2012). \textit{Markov processes, structure and asymptotic behavior: Structure and asymptotic behavior} 184. Springer Science & Business Media.

[51] STOCK, J. H. and WATSON, M. W. (2002). Forecasting using principal components from a large number of predictors. \textit{Journal of the American Statistical Association} 97 1167-1179. · Zbl 1041.62081

[52] STOCK, J. H. and WATSON, M. W. (2016). Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In \textit{Handbook of Macroeconomics}, 2 415-525. Elsevier.

[53] TONG, H. (1990). \textit{Non-linear time series: a dynamical system approach}. Oxford University Press.

[54] TRAPANI, L. (2018). A randomized sequential procedure to determine the number of factors. \textit{Journal of the American Statistical Association} 113 1341-1349. · Zbl 1402.62167

[55] TSAY, R. S. (2005). \textit{Analysis of financial time series} 543. John Wiley & Sons.

[56] TSAY, R. S. and CHEN, R. (2018). \textit{Nonlinear time series analysis} 891. John Wiley & Sons.

[57] WANG, D., LIU, X. and CHEN, R. (2019). Factor models for matrix-valued high-dimensional time series. \textit{Journal of Econometrics} 208 231-248. · Zbl 1452.62684

[58] YE, Z. and WEISS, R. E. (2003). Using the bootstrap to select one of a new class of dimension reduction methods. \textit{Journal of the American Statistical Association} 98 968-979. · Zbl 1045.62034

[59] ZHANG, A. (2019). Cross: Efficient low-rank tensor completion. \textit{The Annals of Statistics} 47 936-964. · Zbl 1416.62298

[60] Zhou, H., Li, L. and Zhu, H. (2013). Tensor regression with applications in neuroimaging data analysis. \textit{Journal of the American Statistical Association} 108 540-552. · Zbl 06195959

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.