Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Stochastic Processes and their Applications xx (xxxx) xxx–xxx

stochastic processes and their applications

www.elsevier.com/locate/spa

# Asymptotics for high-dimensional covariance matrices and quadratic forms with applications to the trace functional and shrinkage

Ansgar Steland[a,*], Rainer von Sachs[b]

[a] *Institute of Statistics, RWTH Aachen University, Wüllnerstr. 3, D-52056 Aachen, Germany*
[b] *Institut de Statistique, Biostatistique et Sciences Actuarielles (ISBA), Université catholique de Louvain, Voie du Roman Pays 20, B-1348 Louvain-la-Neuve, Belgium*

Received 21 September 2016; received in revised form 10 July 2017; accepted 21 October 2017
Available online xxxx

**Abstract**

We establish large sample approximations for an arbitrary number of bilinear forms of the sample variance–covariance matrix of a high-dimensional vector time series using $\ell_1$-bounded and small $\ell_2$-bounded weighting vectors. Estimation of the asymptotic covariance structure is also discussed. The results hold true without any constraint on the dimension, the number of forms and the sample size or their ratios. Concrete and potential applications are widespread and cover high-dimensional data science problems such as tests for large numbers of covariances, sparse portfolio optimization and projections onto sparse principal components or more general spanning sets as frequently considered, e.g. in classification and dictionary learning. As two specific applications of our results, we study in greater detail the asymptotics of the trace functional and shrinkage estimation of covariance matrices. In shrinkage estimation, it turns out that the asymptotics differ for weighting vectors bounded away from orthogonality and nearly orthogonal ones in the sense that their inner product converges to 0.


*MSC:* primary 60F17; secondary 62E20

*Keywords:* Brownian motion; Linear process; Long memory; Strong approximation; Quadratic form; Trace

* Corresponding author.
*E-mail addresses:* steland@stochastik.rwth-aachen.de (A. Steland), rainer.vonsachs@uclouvain.be (R. von Sachs).
*URL:* http://www.stochastik.rwth-aachen.de (R. von Sachs).

https://doi.org/10.1016/j.spa.2017.10.007
0304-4149/

Please cite this article in press as: A. Steland, R. von Sachs, Asymptotics for high-dimensional covariance matrices and quadratic forms with applications to the trace functional and shrinkage, Stochastic Processes and their Applications (2017), https://doi.org/10.1016/j.spa.2017.10.007.

## 1. Introduction

A large number of procedures studied to analyze high- dimensional vector time series of dimension $d_n$ depending on the sample size $n$ rely on projections, e.g. by projecting the observed random vector onto a spanning set of a lower dimensional subspace of dimension $L_n$. Examples include sparse principal component analysis, see e.g. [20], in order to reduce dimensionality of data, sparse portfolio replication and index tracking as studied by [4], or dictionary learning, see [1], where one aims at representing input data by a sparse linear combination of the elements of a dictionary, frequently obtained as the union of several bases and/or historical data.

When studying projections, it is natural to study the associated bilinear form $\mathbf{v}_n' \widehat{\boldsymbol{\Sigma}}_n \mathbf{w}_n$, $\mathbf{v}_n, \mathbf{w}_n \in \mathbb{R}^{d_n}$, representing the dependence structure in terms of the projections' covariances. Here and throughout the paper $\widehat{\boldsymbol{\Sigma}}_n$ is the (uncentered) sample variance–covariance matrix. In order to conduct inference, large sample distributional approximations are needed. For a vector time series model given by correlated linear processes, we established in [17] a strong approximation by a Brownian motion for a single quadratic form, provided the weighting vectors are uniformly bounded in the $\ell_1$-norm. It turned out that the result does not require any condition on the ratio of the dimension and the sample size, contrary to many asymptotic results in high dimensional statistics and probability.

In the present article, we study the more general case of an *increasing* number of quadratic forms as arising when projecting onto a sequence of subspaces whose dimension converges to $\infty$. Noting that the analysis of autocovariances of a stationary linear time series appears as a special case of our approach, there are a few recent results related to our work: [23] established a central limit theorem for a finite number of autocovariances, whereas in [22] the case of long memory series has been studied. [8] has studied the asymptotic theory for detecting a change in mean of a vector time series with growing dimension.

To treat the case of an increasing number of bilinear forms, we consider two related but different frameworks: The first framework uses a sequence of Euclidean spaces $\mathbb{R}^{d_n}$ equipped with the usual Euclidean norm. The second framework embeds those spaces in the sequence space $\ell_2$ equipped with the $\ell_2$-norm. It is shown that, in both frameworks, an increasing number of, say $L_n$, quadratic forms can be approximated by Brownian motions without any constraints on $L_n$, $d_n$ and $n$ apart from $n \to \infty$. One of our main results asserts that, for the assumed time series models, one can define, on a new probability space, equivalent versions and a Gaussian process $\mathcal{G}_n$ taking values in $C([0,1], \mathbb{R}^{L_n})$, such that

$$\sup_{t\in[0,1]} \frac{1}{\sqrt{nL_n}} \left| \left( \mathbf{v}_n^{(j)\prime} (\widehat{\boldsymbol{\Sigma}}_{\lfloor nt \rfloor} - E\widehat{\boldsymbol{\Sigma}}_{\lfloor nt \rfloor}) \mathbf{w}_n^{(j)} \right)_{j=1}^{L_n} - \mathcal{G}_n(t) \right| = o_P(1),$$

as $n \to \infty$, almost surely (a.s.), without any constraints on $L_n$, $d_n$.

We believe that those results have many applications in diverse areas, as indicated above. In this paper we study in some detail two direct applications: The first application considers the trace operator, which equals the trace matrix norm $\|\cdot\|_{tr}$ when applied to covariance matrices. We show that the trace of the sample covariance matrix, appropriately centered, can be approximated by a Brownian motion, a.s. on a new probability space, which also establishes the convergence rate

$$\left| \|\widehat{\boldsymbol{\Sigma}}_n\|_{tr} - \|E\widehat{\boldsymbol{\Sigma}}_n\|_{tr} \right| = O_P(n^{-1/2} d_n).$$

The second application elaborated in this paper is shrinkage estimation of a covariance matrix as studied in depth for i.i.d. sequences of high-dimensional random vectors as well as dependent vector time series, see by [12,13] and [16] amongst others. In order to regularize the sample