V. I. Serdobolskii and A. I. Glusker

Spectra of large sample covariance matrices

Limit spectral theory of sample covariance matrices of increasing dimension was recently used as a base for the development of improved non-degenerating methods of multivariate statistical analysis. We present results of a numerical investigation of fundamental relations of this theory (of the "canonical equations") that thus prove the accuracy of asymptotic relations, find boundaries of the applicability, and the rate of decrease of the remainder terms. The distribution free property of quality functionals for regularized statistical procedures is confirmed experimentally. We show that theoretical upper estimates of the asymptotics remainder terms are 10–100 times overstated.

For the last decade, an extensive limit spectral theory of random matrices created by V.L.Girko and other authors [1,2] was successfully applied to sample covariance matrices [3–11]. Using asymptotic relations of spectral theory it was possible to construct a number of improved and unimprovable statistical procedures of multivariate analysis [9]. New procedures have a number of substantial advantages over traditional consistent ones: they do not degenerate for any data, are applicable and approximately optimal property is guaranteed independently on distributions. Most of these results were obtained in the asymptotic approach in which the observation dimension $n \to \infty$ along with sample sizes N so that $n/N \to y > 0$. This asymptotics serves as a tool of isolating principle parts of quality functions, constructing their estimators, and the solution of extremal problems. The problem of practical applicability of the improved procedures is now reduced to the estimation of the magnitude of remainder terms. Some theoretical upper estimates of the remainder terms were obtained in [7] and [9]. However, these estimates seem to be too restrictive for applications. The question of practical advantages of asymptotically improved and unimprovable procedures remains open.

In this paper we offer an experimental numerical investigation of spectral functions of large sample covariance matrices and the comparison with limit relations of spectral theory.

1 Theoretical Results

The remarkable feature of the spectral theory of random matrices of increasing dimension is its independence on distributions. In [4] and [9] it is shown that applicability of asymptotic formulas of this theory is provided by magnitudes of the following two parameters. Let $\mathbf{x}$ be an observation vector, $\mathcal{S}$ be a population for which $\mathbf{E}\,\mathbf{x} = 0$. Denote

$$M = \sup_{|\mathbf{e}|=1} \mathbf{E}(\mathbf{e}^T\mathbf{x})^4 > 0, \tag{1}$$

where $\mathbf{e}$ are non-random unity vectors;

$$\gamma = \sup_{\|\Omega\|\le 1} \text{ var } (\mathbf{x}^T\Omega\mathbf{x}/n)/M, \tag{2}$$

where Ω are non-random symmetric positive semidefinite matrices with spectral norms not greater 1 (we will use only spectral norms of matrices).

Denote $\Sigma = \text{cov}(\mathbf{x}, \mathbf{x})$. For an non-degenerate normal distribution $\mathbf{N}(0, \Sigma)$**,** the values $M = 3\|\Sigma\|^2$, $\gamma = 2/3\ n^{-2}\text{tr }\Sigma^2$. If $\Sigma = I$ $M = 3$, $\gamma = 2/3n$. Let $\mathcal{X} = \{\mathbf{x}_\text{m}\}$ be a sample of size N from population $\mathcal{S}$. Let

$$\bar{\mathbf{x}} = N^{-1}\sum_{m=1}^{N}\mathbf{x}_m, \quad C = N^{-1}\sum_{m=1}^{N}(\mathbf{x}_m - \bar{\mathbf{x}})(\mathbf{x}_m - \bar{\mathbf{x}})^T.$$

Here C is a sample covariance matrix (a biased estimator of Σ).

To be concise, denote $y = n/N$, $\tau = \sqrt{M}t$,

$$H = H(t) = (I + tC)^{-1}, \quad h_n(t) = n^{-1}\text{tr } H(t), \quad s_n(t) = 1 - y + yh_n(t).$$

The "canonic" relation between spectral functions of the matrices Σ and C is presented by the following statement.

Theorem 1 (corollary of [4]). *For any populations with $M > 0$, for $t \ge 0$,*

$$h_n(t) = n^{-1}\,\text{tr}(I + t\widehat{s}(t)\Sigma)^{-1} + \omega(t), \tag{3}$$

$$\mathbf{E}\,H(t) = (I + ts_n(t)\Sigma)^{-1} + \Omega, \tag{4}$$

and

$$\text{var } \left(n^{-1}H(t)\right) \le a\tau^2/N,$$

where $\omega(t) = \|\Omega\| \le a\tau \ \max\,(1, y)[\tau\sqrt{\delta} + (1 + \tau^2)/\sqrt{N}], \quad \delta = 2y^2(\gamma + \tau^2/N)$,

and a is a numeric constants.

In Chapter 1 of [2] , the expectation of $n^{-1}\text{tr } H(t)$ was found more accurately, and in [5] it was shown that the variance of this function is $O(n^{-1}N^{-1}$. Using these results, we can offer a stronger estimator for the case of $\Sigma = I$:

$$\mathbf{E}\,|\omega(t)| \le \omega_1(t) \stackrel{\text{def}}{=} 2yt^2(1 + ty)/\sqrt{nN} + 3t/N$$

The spectral equations (3) and (4) relate spectra of the observed covariance matrices C to spectra of unknown matrices Σ.

The problem of recovering spectral functions for the matrices Σ by matrices C can be solved in the form of limit relations. To pass to the limit, we consider a sequence of problems of the statistical analysis

$$\mathcal{P} = \{(\Pi, \Sigma, \text{N}, \mathcal{X}, \bar{\mathbf{x}}, \text{C})_n\},\ n = 1, 2, \ldots, \tag{5}$$

where (indexes n are omitted) Π is an n-dimensional population with $\Sigma = \text{cov}(\mathbf{x}, \mathbf{x})$, $\mathcal{X}$ is a sample of size N.

Theorem 2 (a corollary from [3], [4] and [9]). *Let the moments* (1) *exist and be uniformly bounded in* $\mathcal{P}$*; as* $n \to \infty$ *let* $n/N \to \lambda > 0$, $\gamma \to 0$, *and for almost all* $u \geq 0$ *the limit exists*

$$F_0(u) \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \text{ind}(\lambda_i^0 \leq u), \tag{6}$$

where $\lambda_1^0, \ldots, \lambda_n^0$ *are the eigenvalues of* Σ.

Then for each $t \geq 0$

1/ there exists the limit

$$h(t) = \lim_{n\to\infty} \mathbf{E}\, n^{-1}\text{tr}\ H(t) = \int (1 + ts(t)u)^{-1} dF_0(u), \tag{7}$$

where $s(t) = 1 - \lambda + \lambda h(t)$;

2/ the variance $\quad \text{var}\ \ (n^{-1}\,\text{tr}\, H(t)) \to 0 \quad n \to \infty$.

The "canonical"integral equation (7) presents the main result of the limit spectral theory of sample covariance matrices of increasing dimension.

Consider an empiric distribution function for eigenvalues λ_i of matrices C

$$F_n(u) = n^{-1} \sum_{i=1}^{n} \text{ind}\ (\lambda_i \leq u), \quad u \geq 0. \tag{8}$$

Theorem 3 (see [4]). *Let assumptions of Theorem* 2 *hold and, in addition,* $\lambda > 0$ *and all eigenvalues of all matrices* Σ *in* $\mathcal{P}$ *exceed* $c_1 > 0$, *where* c_1 *does not depend on* n.

Then the function $h(z)$ *allows an analytical continuation to the region of complex* z *and satisfies the Gölder condition for all* z*; the weak convergence in probability holds* $F_n(u) \to F(u)$, $\quad u \geq 0$*; and for all* z, *except* $z < 0$,

$$h(z) = \int (1 - zu)^{-1} dF(u).$$

In [6] it is shown that as $n \to \infty$ all eigenvalues of matrices C lay on the limit support almost surely.

Consider a special case: let $\Sigma = I$, $n = 1, 2, \ldots$ (see [9], Chapter 2). In this case, the function $h(z)$ satisfies the quadratic equation $h(z) - 1 = zh(z)s(z)$, and the spectral density of C is

$$F'(u) = \sqrt{(u_2 - u)(u - u_1)}/(2\pi\lambda u), \tag{9}$$

where $u_1 = (1 - \sqrt{\lambda})^2$, $u_2 = (1 + \sqrt{\lambda})^2$ for $u_1 \leq u \leq u_2$ The function

$$h(t) = 2\ \left(\sqrt{(1 + (1 - \lambda)t)^2 + 4t\lambda} + 1 + (1 - \lambda)t\right)^{-1} \tag{10}$$

Suppose, additionally, that there exists an $\varepsilon > 0$ such that the quantity

$$\sup_{|\mathbf{e}|=1} \mathbf{E}(\mathbf{e}^T\mathbf{x})^{4+\varepsilon}$$

is uniformly bounded in $\mathcal{P}$ and the condition (1.3) from [3] holds. Then Theorem 11.1 from [3] states that as $n \to \infty$ the minimum and maximum eigenvalues of the matrices C converge in probability to the magnitudes α_1 and α_2 such that

$$\alpha_i = \operatorname{plim}\left(1 - \frac{y}{n}\sum_{k=1}^{n}\frac{\lambda_k}{\lambda_k - x_i}\right), \quad i = 1, 2,$$

where $x = x_1 \;\; x = x_2$ are the minimal and maximal real roots of the equation

$$\frac{y}{n}\sum_{k=1}^{n}\frac{\lambda_k^2}{(\lambda_k - x)^2} = 1.$$

2 Numeric Experiments

Numeric experiments were performed with the purpose

- to investigate the relation between spectral functions of true covariance matrices Σ and sample covariance matrices C, given by Equations (3) and (7);
- to study the convergence of the empirical distribution function of eigenvalues of C for large n;
- to recover spectra of unknown true covariance matrices by observations using Girko's G-estimator and to estimate the accuracy of the inverse problem solution;
- to investigate boundaries of spectra of matrices C and to compare these with the theory [4];
- to study the dependence of the remainder terms $\omega(t)$ in (3) and (5) on the parameters t, M, $y = n/N$ and γ, and to compare with the upper estimates;
- and to estimate experimentally the boundaries of the applicability of the asymptotic equation (3) to distributions different from normal.

Given a fixed distribution law, n, and N, samples were generated and used for calculation of sample means, sample covariance matrices, and spectral functions $h_n(t)$, $H(t)$, and $F_n(u)$ (denote $\widehat{h}(t) = h_n(t), \;\; \widehat{F}(u) = F_n(u)$ and their averages over s experiments by $\langle\widehat{h}(t)\rangle$ and $\langle\widehat{F}(u)\rangle$). These values were sharpened in a series of s experiments. The random inaccuracy $\omega(t)$ of the equation (3) was also calculated in a series of s experiments (an estimator of $\mathbf{E}\;\omega(t)$), and its mean square deviation was found (an estimator of the variance var $\omega(t)$). These characteristics are compared in tables with theoretical upper estimates of the asymptotic formulas inaccuracy.

2.1 Empirical spectral function $\widehat{h}(t)$. Distribution $x \sim N(0, I)$.

In Tables 1–4 the first two lines present the values of the function $\widehat{h}(t)$ for two independent experiments, the third line contains an estimator of the expectation along with the mean square deviation (in a series of $s = 100$ experiments). In the last line, the theoretical function $h(t)$ is shown by (9). The analytical dependence of the deviations of the form $|\langle\widehat{h}(t)\rangle - h(t)| = kN^{-b}$ was fitted with the minimum square method: it is shown under the tables with the mean square error.

Table 1.
Table 2.
Table 3.
Table 4.

From Tables 1–4 one can see the systematic decrease of $\langle\widehat{h}(t)\rangle$ with growing N for fixed $y = n/N$ that can be explained by the increase of accuracy in the resolvent denominator. Values of $\langle\widehat{h}(t)\rangle$ tend to theoretical functions $h(t)$ as N increases, and the difference decreases approximately by the law N^{-1}. The scatter of empiric $\langle\widehat{h}(t)\rangle$ in two experiments is covered by 2.5 σ, where σ^2 is the sample variance.

2.2 Accuracy of the basic spectral equation. Distribution $\mathbf{x} \sim N(0, I)$

In Tables 5-8 the experimental inaccuracy $\omega(t)$ for the equation (3) is presented. The value $\widehat{s}(t)$ was calculated by the formula $\widehat{s}(t) = 1 - y + y\widehat{h}(t)$, where $\widehat{h}(t)$ was taken from Tables 1–4.

As in the former tables, the first two lines present two independent experimental values of $\widehat{\omega}(t)$, the next line shows sample mean and the mean square deviation in a series of $s = 100$ experiments. The fourth line contains theoretical upper inaccuracy $\omega(t)$ in (3) for $a = 1$ (this estimator is distribution free). The 5th line presents the minimum value of a, for which (3) holds with $\omega(t)$ substituted by its expectation; in the 6th line, the refined theoretical upper estimate $\omega_1(t)$ is shown (valid for normal distribution).

Using $\widehat{\omega}(t)$ the analytical dependence $\omega(t) = kN^{-c}$ on N was constructed by the minimum square method for the expectation and the square scatter. It is shown under the tables.

Table 5.
Table 6.
Table 7.
Table 8.

From Tables 5-8 one can see that the average inaccuracy (in the series of 100 experiments) of the equation (3) $\widehat{\omega}(t)$ decreases approximately by the law N^{-1} as $n = N$ increases.

Theoretical upper estimate $\omega(t)$ decreases as $N^{-1/2}$, while the more refined estimator $\omega_1(t)$ (valid for normal distributions) decreases as N^{-1}. For $n = N$, the upper estimate $\omega(t)$ in all cases is comparable with 1 or more and has no sense as much the inaccuracy $\omega(t)$ is not greater 1 by definition. The upper estimate $\omega_1(t)$ proves to be ten times overstated as compared with the experimental value. For $n = N$, the mean square deviation of $\omega(t)$ decreases by the law N^{-1} that corresponds to the estimator for the variance of $\widehat{h}(t)$ for normal distribution found in [5], while (3) guarantees only the law $N^{-1/2}$. The estimator of the variance of $\widehat{h}(t)$ in [5] proves to be overstated by 5 times as compared with the empiric mean square deviation of $\omega(t)$ for $t = 1$ and $n = N = 50$. The disagreement with the experiment increases with the increase of t.

In Fig.1 for $n = N = 50$, plots of functions $\widehat{h}(t)$ (solid line) and $h(t)$ (discontinuous line) are presented together with their difference (dotted line).

Fig. 1.

2.3 Spectral function $\widehat{h}(t)$ for binomial and normal populations)

In Tables 9–10 the estimators $\langle\widehat{h}(t)\rangle$ of $h(t)$ expectation are presented calculated in a series of $s = 100$ experiments together with the mean square deviation $\sigma(t)$ of $\widehat{h}(t)$. Spectral functions are compared for normal distribution $N(0, I)$ and discrete binomial law $B(1, p)$ (in which each component of $\mathbf{x}$ takes on values $a = -\sqrt{p^{-1}(1-p)}$ or $b = \sqrt{p(1-p)^{-1}}$ $(p > 0)$ with the probabilities $\mathbf{P}(a) = p$, $\mathbf{P}(b) = 1 - p$ so that $\mathbf{E}\,\mathbf{x} = 0$ and $\mathbf{E}\,\mathbf{x}^2 = 1$.

In the last line, limit values $h(t)$ are shown calculated by (7).

Table 9.

Table 10.

One can see a good agreement of experimental results for these two distributions and a good fit to theoretical limits as well as a small scatter of estimation. Theoretically expected mean square scatter can be measured by the quantity $(nNs)^{-1/2} = 0.002$. This fact demonstrates well the insensibility of spectral functions of sample covariance matrices to distributions when dimension is high $(n = 50)$.

2.4 Function $\widehat{h}(t)$ for non-standard normal distribution

In Table 11 the results of numerical modeling are presented (the estimator of the expectation and its mean square deviation calculated in a series of experiments) for normal distribution

with the covariance matrix

$$\Sigma = \begin{pmatrix} 2 & 1 & -1 & \dots & 1 & -1 \\ -1 & 2 & -1 & \dots & 1 & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & 1 & -1 & \dots & 2 & -1 \\ -1 & 1 & -1 & \dots & 1 & 2 \end{pmatrix},$$

which corresponds to observations with the correlation coefficient $\pm 1/2$ and variance 2. The eigenvalues of this matrix include (for even n) $n/2$ units and $n/2$ of 3s. In the lowest line, the limit function $h(t)$ is shown. This function was calculated by (7) as a root of the equation

$$2h = (1+ts)^{-1} + (1+3ts)^{-1},\ h = h(t),\ s = s(t).$$

Table 11.

One can see a good agreement between the experimental and theoretical values of the function $\widehat{h}(t)$.

2.5 Function $\widehat{h}(t)$ and the inaccuracy $\widehat{\omega}(t)$ for different coupling of variables

Tables 12-14 present results of numeric simulation of the function $\widehat{h}(t)$ and the inaccuracy $\widehat{\omega}(t)$ (estimator of the expectation and its mean square scatter in $s = 100$ experiments) for normal population with the covariance matrix

$$\Sigma = \begin{pmatrix} 1 & r & \dots & r & r \\ r & 1 & \dots & r & r \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ r & r & \dots & 1 & r \\ r & r & \dots & r & 1 \end{pmatrix} \tag{11}$$

for different $r < 1$. This matrix has a spectrum with a single eigenvalue $\lambda_1 = 1 - r + rn$ and $n-1$ eigenvalues with $\lambda = 1 - r$. For large n, the parameter τ in (3) approximately equals $\sqrt{3rn}$. Theoretical limit value of this function can be calculated from (7) and is

$$h(t) = 2\left(\sqrt{(1+t(1-y)(1-r))^2 + 4ty(1-r)} + 1 + t(1-y)(1-r)\right)^{-1}$$

Table 12.
Table 13.
Table 14.

These tables show a good fit between experimental average values of $\widehat{h}(t)$ and limit values calculated from the basic equation (7). The good fit keeps on even for $r = 0.9$. Theoretical

upper estimates of $\omega(t)$ are strongly overstated. For large matrices of the form (11), the fourth maximum moment $M \approx 3r^2n^2$, and theoretical upper estimate increases n^2 times. These estimates are proportional to powers of M, whereas Tables 12–14 show the linear increase of $\omega(t)$ with r or $\sqrt{r}$. Thus, the basic limit spectral equation (7) remains also valid for strongly coupled variables (up to $r = 0.9$).

2.6 Spectra of sample covariance matrices. Distribution $\mathbf{x} \sim N(0, I)$

Sample covariance matrices C were simulated. Table 15 shows maximal eigenvalues of C together with mean square deviations and the deviation of the averaged value from the theoretical limit value $\lambda_{max} = 4$. Below the table, the empirical regularity is shown obtained by the minimum square method.

Table 15.

In Table 16, the sample mean and the mean square deviation is presented for the uniform norm of the difference $\|\widehat{F} - F\|$ of functions $\widehat{F}(u)$ and $F(u)$, where $\widehat{F}(u)$ was calculated by (8), and $F'(u)$ is the theoretical density. Table 17 presents sample mean of the function $\widehat{F}(1)$ and its mean square deviation calculated in a series of experiments. Below, the empiric formulas are shown fitting the data from Tables 16 and 17.

Table 16.

Table 17.

One can see that the disagreement with the theoretical limit formula (7) decreases proportionally to $N^{-0.8}$ and $N^{-1.0}$ and is approximately 0.04 for the uniform norm of the difference, and 0.007 for the function at the point $u = 1$. This fact indicates the lessening of the agreement at the endpoints of the spectra.

For the illustration, we present a plot showing theoretical and experimemtal distribution functions $F(u)$ and the corresponding density $F'(u)$ for $n = N = 75$ (Fig. 2.).

Fig. 2

Solid lines denote the empiric distribution function $\widehat{F}(u)$ (decreasing) and the smoothed density $\widehat{F}'(u)$ (increasing); discontinuous lines denote limit functions $F(u)$ and $F'(u)$.

2.7 Girko's G_2-estimator

V. L. Girko ([2], Chapter 5) suggested an estimator for the spectral function of unknown true covariance matrices Σ. To estimate $\eta(t) = n^{-1}\mathrm{tr}\ (I + t\Sigma)^{-1}$, he proposed the statistics $\widehat{\eta}(t) = h(\theta)$, where θ is a root of the equation $t = \theta(1 - y + y\widehat{h}(\theta))$, and $y = n/N$. In [3] it is proved that under some assumptions, this equation is always solvable for $t > 0$ and defines a function that converges as $n \to \infty$ and $n/N \to \lambda$ almost surely to $\eta(t)$. Experimental investigation of this estimator is shown in Table 18. The empirical function

$\widehat{\eta}(t)$ was calculated by spectra of matrices C by averaging over $s = 100$ experiments along with its mean square deviation.

For the comparison, we also show the theoretical value $\eta(t) = (1 + t)^{-1}$. In the lower line, the empirical laws are shown for the the deviation $\langle \eta(t) \rangle$ from $\eta(t)$.

Table 18.

This table demonstrates the convergence of G-estimator by the law $O(N^{-1})$.

3 Conclusions

Our numerical experiments show first that spectral functions of sample covariance matrices of large dimension converge as $n \to \infty$ and $n/N \to \lambda$, where n is the dimension of observations, and N are sample sizes. The empiric variance of the function $h_n(t) = n^{-1}\mathrm{tr}\ (I+tC)^{-1}$ decreases by the law near to $n^{-1}N^{-1}$ as predicted by the theorical estimate [5]. The theoretical upper estimate for the variance of $h_n(t)$ by Theorem 1 shows only the decrease by the law N^{-1} that seems to be related to taking no account of the independence of the observation vector components. The theoretical coefficient of $n^{-1}N^{-1}$ for the asymptotic variance proves to be excessive by factor 5–10 . In all experiments we observed the convergence of $\widehat{h}(t)$ to the theoretical value $h(t)$ defined by the canonic spectral equation (7). The difference between $h(t)$ and averaged values of $\widehat{h}(t)$ decreases approximately as N^{-1}. The theoretical upper estimates of the inaccuracy of asymptotic formulas found in [7] and [9] prove to be substantially overstated (by hundreds times). A more precise upper estimate of this inaccuracy in [5] proves to be overstated by 2–4 times.

The applicability of all asymptotic theory [4–10] is restricted by the magnitude of the maximal invariant fourth momentum (1) that increases as n^2 as the dimension $n \to \infty$ if variables are strongly coupled and all correlation coefficients equal 1. The dependence of the accuracy of the asymptotic equation (3) on variables coupling (tables 12–14) show that the theoretical requirements to small coupling seem too stringent. The empirical law of the dependence of inaccuracy on the correlation coefficient r does not agree with theoretical laws for upper estimates of the remainder terms $\omega(t)$ and $\omega_1(t)$. Experiments show that, for normal populations, the equation (7) holds well even for $r \approx 0.8$ when $n = N = 50$. Thus, methods of estimation of the remainder terms developed in [7] and [9] provide only weak estimates, that require sharpening and revision. It remains not quite clear what minimal restrictions on the dependence of variables are yet necessary for the canonic equations [7] to be accurate.

Our experiments allow to make a general conclusion that the basic canonic spectral equations for large sample covariance matrices proves to be accurate even for not large n and N: the average inaccuracy in different experiments decreases from 0.05 for $n = N = 10$ to 0.01 for $n = N = 50$. This result also keeps for non-standard normal distribution and for other distributions, and in particular, for discrete distributions and remains true for some cases of strongly coupled variables.

In [10] a "Normal evaluation principle" was offered stating that standard quality functionals of regularized multivariate statistical procedures are only weakly dependent on moments of variables higher than the second, and thus, are approximately the same as for normal distributions. This property is a consequence of the corresponding properties of spectral functions of sample covariance matrices that are confirmed experimentally.

Our results substantiate the asymptotical approach of [7] and [9] and show that the improved methods presented in these investigations are expected to have the accuracy well acceptable for applied statistical investigations even for not large parameters n and N (possibly, even for $n = N = 5 - 10$).

References

1. *V. L. Girko*, Random Matrices (in Russian), Kiev, Kiev State University, 1975.

2. *V. L. Girko*, Spectral Theory of Random Matrices (in Russian), Moscow, "Nauka", 1988.

3. *V. Girko*, Statistical Analysis of Observations of Increasing Dimension, Kluwer Academic Publishers, Dordrecht, 1995

4. *V. I. Serdobolskii*, Spectral properties of sample covariance matrices, Theory Probab. Appl., 40, 1995.

5. *T.A.Baranova and I.F.Sentiureva*, On the accuracy of asymptotic expressions for spectral functions and the resolvent of sample covariance matrices (in Russian), in: Random Analysis, Moscow, Moscow State University, 1987, 17–24.

6. *Z. D. Bai and J. W. Silverstein*, No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. The Annals of Probab., vol. 26, 316–345.

7. *V. I. Serdobolskii*, Theory of essentially multivariate statistical analysis, Russian Math. Surveys, 54, 2, 1999.

8. *V. I. Serdobolskii*, Application aspects of G-analysis, Random Operators and Stochastic Equations, 1999, vol. 6, no. 4, 339–352

9. *V. Serdobolskii*, Multivariate Statistical Analysis. A High-Dimensional Approach, Kluwer Academic Publishers, Dordrecht, 2000

10. *V. I. Serdobolskii*, Normal model for distribution-free multivariate analysis, Statistics and Probability Letters, 48, 2000, 353–360.

Table 1: Spectral function $\widehat{h}(t)$: $\mathbf{x} \sim N(0, I)$, $t = 1$, $y = 1$, $s = 100$

	$N = 2$	$N = 4$	$N = 10$	$N = 20$	$N = 50$
$\widehat{h}(t) : 1$	0.806	0.711	0.634	0.652	0.622
$\widehat{h}(t) : 2$	0.820	0.714	0.630	0.612	0.619
$\langle \widehat{h}(t) \rangle$	0.794 ± 0.107	0.699 ± 0.068	0.657 ± 0.025	0.639 ± 0.014	0.626 ± 0.005
$h(t)$	0.618	0.618	0.618	0.618	0.618

Empirical law: $\langle \widehat{h}(t) \rangle - h(t) \approx 0.32 N^{-0.93} \pm 0.23 N^{-0.96}$

Table 2: Spectral function $\widehat{h}(t)$: $\mathbf{x} \sim N(0, I)$, $t = 2$, $y = 1$, $s = 100$

	$N = 2$	$N = 4$	$N = 10$	$N = 20$	$N = 50$
$\widehat{h}(t) : 1$	0.750	0.615	0.591	0.504	0.510
$\widehat{h}(t) : 2$	0.874	0.629	0.515	0.537	0.508
$\langle \widehat{h}(t) \rangle$	0.700 ± 0.103	0.632 ± 0.064	0.544 ± 0.028	0.521 ± 0.015	0.508 ± 0.006
$h(t)$	0.500	0.500	0.500	0.500	0.500

Empirical law: $\langle \widehat{h}(t) \rangle - h(t) \approx 0.46 N^{-1.0} \pm 0.21 N^{-0.90}$

Table 3: Spectral function $\widehat{h}(t)$: $\mathbf{x} \sim N(0, I)$, $t = 1$, $y = 1/2$, $s = 100$

	$N = 2$	$N = 4$	$N = 10$	$N = 20$	$N = 50$
$\widehat{h}(t) : 1$	0.493	0.815	0.555	0.580	0.573
$\widehat{h}(t) : 2$	0.537	0.815	0.585	0.578	0.566
$\langle \widehat{h}(t) \rangle$	0.740 ± 0.201	0.676 ± 0.120	0.588 ± 0.048	0.581 ± 0.022	0.568 ± 0.010
$h(t)$	0.562	0.562	0.562	0.562	0.562

Empirical law: $\langle \widehat{h}(t) \rangle - h(t) \approx 0.42 N^{-1.1} \pm 0.43 N^{-0.96}$

Table 4: Spectral function $\widehat{h}(t)$: $\mathbf{x} \sim N(0, I)$, $t = 2$, $y = 1/2$, $s = 100$

	$N = 2$	$N = 4$	$N = 10$	$N = 20$	$N = 50$
$\widehat{h}(t) : 1$	0.392	0.358	0.504	0.446	0.417
$\widehat{h}(t) : 2$	0.170	0.736	0.369	0.422	0.422
$\langle \widehat{h}(t) \rangle$	0.708 ± 0.233	0.530 ± 0.132	0.459 ± 0.059	0.435 ± 0.026	0.424 ± 0.010
$h(t)$	0.414	0.414	0.414	0.414	0.414

Empirical law: $\langle \widehat{h}(t) \rangle - h(t) \approx 0.52 N^{-1.0} \pm 0.50 N^{-0.98}$

Table 5: The remainder term $\omega(t)$ in Equation (**??**): $\mathbf{x} \sim N(0, I)$, $y = 1$, $t = 1$, $s = 100$

	$N = 4$	$N = 6$	$N = 10$	$N = 20$	$N = 50$
$\omega(t) : 1$	0.120	0.0731	0.0570	0.0303	0.0161
$\omega(t) : 2$	0.115	0.111	0.0637	0.0308	0.0110
$\langle \omega(t) \rangle$	0.117 ± 0.058	0.0815 ± 0.0373	0.0526 ± 0.0284	0.0254 ± 0.0141	0.00959 ± 0.00525
$\omega(t)$	7.53	6.15	4.76	3.37	2.13
a	0.0155	0.0133	0.0111	0.00755	0.0045
$\omega_1(t)$	1.75	1.17	0.70	0.350	0.14

Empirical law: $\langle \omega(t) \rangle \approx 0.62 N^{-1.2} \pm 0.22 N^{-0.93}$

Table 6: The remainder term $\omega(t)$ in Equation (**??**): $\mathbf{x} \sim N(0, I)$, $y = 1$, $t = 2$, $s = 100$

	$N = 4$	$N = 6$	$N = 10$	$N = 20$	$N = 50$
$\omega(t) : 1$	0.172	0.147	0.0347	0.0668	0.0128
$\omega(t) : 2$	0.193	0.0463	0.0880	0.0469	0.0113
$\langle \omega(t) \rangle$	0.167 ± 0.064	0.107 ± 0.047	0.0684 ± 0.0329	0.0309 ± 0.0162	0.0129 ± 0.0057
$\omega(t)$	52.7	43.0	33.3	23.6	14.9
a	0.00318	0.00248	0.00205	0.00131	0.000862
$\omega_1(t)$	7.50	5.00	3.00	1.50	0.60

Empirical law: $\langle \omega(t) \rangle \approx 0.65 N^{-1.0} \pm 0.24 N^{-0.92}$

Table 7: The remainder term $\omega(t)$ in Equation (**??**): $\mathbf{x} \sim N(0, I)$, $y = 1/2$, $t = 1$, $s = 100$

	$N = 4$	$N = 6$	$N = 10$	$N = 20$	$N = 50$
$\omega(t) : 1$	-0.151	0.168	-0.0252	0.0561	0.0154
$\omega(t) : 2$	0.0928	0.158	0.0253	0.0102	0.0153
$\langle\omega(t)\rangle$	0.124 ± 0.097	0.0784 ± 0.0746	0.0416 ± 0.0442	0.0216 ± 0.0237	0.00970 ± 0.00857
$\omega(t)$	2.84	2.32	1.79	1.27	0.802
a	0.0438	0.0338	0.0232	0.017	0.0121
$\omega_1(t)$	1.28	0.85	0.51	0.26	0.10

Empirical law: $\langle\omega(t)\rangle \approx 0.45N^{-1.0} \pm 0.44N^{-1.0}$

Table 8: The remainder term $\omega(t)$ in Equation (**??**): $\mathbf{x} \sim N(0, I)$, $y = 1/2$, $t = 2$, $s = 100$

	$N = 4$	$N = 6$	$N = 10$	$N = 20$	$N = 50$
$\omega(t) : 1$	0.285	0.321	0.0707	-0.0126	0.0182
$\omega(t) : 2$	0.0647	0.0780	0.0906	0.0332	0.0175
$\langle\omega(t)\rangle$	0.169 ± 0.127	0.0897 ± 0.0843	0.0558 ± 0.0456	0.0284 ± 0.0241	0.0121 ± 0.0109
$\omega(t)$	19.0	15.5	12.0	8.50	5.38
a	0.00891	0.00578	0.00464	0.00334	0.00224
$\omega_1(t)$	4.33	2.89	1.73	0.87	0.35

Empirical law: $\langle\omega(t)\rangle \approx 0.58N^{-1.0} \pm 0.49N^{-1.0}$

Table 9: Spectral function $\langle\widehat{h}(t)\rangle \pm \sigma(t)$: $y = 1$, $N = 50$, $s = 100$

	$t = 0.1$	$t = 0.25$	$t = 0.5$	$t = 1$	$t = 2$
N	0.917 ± 0.002	0.832 ± 0.004	0.736 ± 0.005	0.625 ± 0.006	0.509 ± 0.007
$B(0.3)$	0.918 ± 0.001	0.831 ± 0.002	0.736 ± 0.003	0.624 ± 0.005	0.507 ± 0.005
$B(0.5)$	0.918 ± 0.000	0.831 ± 0.001	0.736 ± 0.002	0.623 ± 0.003	0.505 ± 0.004
$B(0.7)$	0.918 ± 0.001	0.831 ± 0.002	0.736 ± 0.003	0.624 ± 0.004	0.507 ± 0.006
$h(t)$	0.916	0.828	0.732	0.618	0.500

Table 10: Spectral function $\langle\widehat{h}(t)\rangle \pm \sigma(t)$: $y = 1/2$, $N = 50$, $s = 100$

	$t = 0.1$	$t = 0.25$	$t = 0.5$	$t = 1$	$t = 2$
N	0.914 ± 0.003	0.819 ± 0.005	0.709 ± 0.006	0.572 ± 0.009	0.423 ± 0.011
$B(0.3)$	0.914 ± 0.002	0.818 ± 0.004	0.705 ± 0.005	0.566 ± 0.007	0.421 ± 0.007
$B(0.5)$	0.914 ± 0.000	0.818 ± 0.001	0.705 ± 0.003	0.565 ± 0.004	0.418 ± 0.005
$B(0.7)$	0.914 ± 0.002	0.818 ± 0.004	0.705 ± 0.005	0.567 ± 0.007	0.420 ± 0.007
$h(t)$	0.913	0.815	0.702	0.562	0.414

Table 11: Spectral function $\widehat{h}(t)$: $N(0, \Sigma)$, $\Sigma \neq I$, $N = 20$, $s = 100$

	$y = 1$ $t = 0.5$	$y = 1$ $t = 1$	$y = 1$ $t = 2$	$y = 1/2$ $t = 0.5$	$y = 1/2$ $t = 1$	$y = 1/2$ $t = 2$
$\langle\widehat{h}(t)\rangle$	0.65 ± 0.01	0.54 ± 0.01	0.44 ± 0.02	0.60 ± 0.02	0.47 ± 0.03	0.34 ± 0.03
$h(t)$	0.64	0.52	0.42	0.59	0.45	0.32

Table 12: Spectral function $\widehat{h}(0.5)$: $N(0, \Sigma)$, $N = 10$, $s = 100$

	$y = 1$			$y = 0.5$		
	$\langle\widehat{h}(t)\rangle$	$h(t)$	$\langle\widehat{\omega}(t)\rangle$	$\langle\widehat{h}(t)\rangle$	$h(t)$	$\langle\widehat{\omega}(t)\rangle$
$r = 0$	0.755 ± 0.023	0.732	0.03 ± 0.02	0.723 ± 0.039	0.702	0.03 ± 0.04
$r = 0.2$	0.769 ± 0.022	0.766	0.05 ± 0.02	0.766 ± 0.035	0.742	0.07 ± 0.04
$r = 0.4$	0.791 ± 0.019	0.805	0.07 ± 0.02	0.759 ± 0.033	0.788	0.07 ± 0.04
$r = 0.6$	0.820 ± 0.017	0.854	0.11 ± 0.02	0.777 ± 0.031	0.844	0.09 ± 0.03
$r = 0.8$	0.861 ± 0.012	0.916	0.16 ± 0.01	0.819 ± 0.023	0.913	0.13 ± 0.02
$r = 0.9$	0.888 ± 0.009	0.954	0.20 ± 0.01	0.838 ± 0.019	0.953	0.16 ± 0.02

Table 13: Spectral function $\widehat{h}(1)$: $N(0,\Sigma)$, $N = 10$, $s = 100$

	$y = 1$			$y = 0.5$		
	$\langle\widehat{h}(t)\rangle$	$h(t)$	$\langle\widehat{\omega}(t)\rangle$	$\langle\widehat{h}(t)\rangle$	$h(t)$	$\langle\widehat{\omega}(t)\rangle$
$r = 0$	0.652 ± 0.030	0.618	0.05 ± 0.03	0.595 ± 0.044	0.562	0.05 ± 0.04
$r = 0.2$	0.667 ± 0.028	0.656	0.06 ± 0.03	0.654 ± 0.043	0.608	0.11 ± 0.04
$r = 0.4$	0.696 ± 0.023	0.703	0.11 ± 0.02	0.649 ± 0.037	0.667	0.10 ± 0.04
$r = 0.6$	0.743 ± 0.019	0.766	0.16 ± 0.02	0.677 ± 0.042	0.742	0.14 ± 0.04
$r = 0.8$	0.805 ± 0.016	0.854	0.25 ± 0.02	0.746 ± 0.026	0.844	0.21 ± 0.03
$r = 0.9$	0.850 ± 0.010	0.916	0.31 ± 0.01	0.786 ± 0.020	0.913	0.26 ± 0.02

Table 14: Spectraal function $\widehat{h}(2)$: $N(0,\Sigma)$, $N = 10$, $s = 100$

	$y = 1$			$y = 0.5$		
	$\langle\widehat{h}(t)\rangle$	$h(t)$	$\langle\widehat{\omega}(t)\rangle$	$\langle\widehat{h}(t)\rangle$	$h(t)$	$\langle\widehat{\omega}(t)\rangle$
$r = 0$	0.543 ± 0.026	0.500	0.07 ± 0.03	0.457 ± 0.044	0.414	0.06 ± 0.05
$r = 0.2$	0.558 ± 0.030	0.538	0.09 ± 0.03	0.541 ± 0.040	0.461	0.15 ± 0.04
$r = 0.4$	0.593 ± 0.030	0.587	0.14 ± 0.03	0.533 ± 0.039	0.523	0.14 ± 0.04
$r = 0.6$	0.652 ± 0.025	0.656	0.22 ± 0.03	0.559 ± 0.036	0.608	0.17 ± 0.04
$r = 0.8$	0.732 ± 0.019	0.766	0.33 ± 0.02	0.652 ± 0.034	0.742	0.27 ± 0.03
$r = 0.9$	0.801 ± 0.016	0.854	0.42 ± 0.02	0.719 ± 0.024	0.844	0.36 ± 0.03

Table 15: Maximal eigenvalues of the matrix C: $y = 1$, $N(0, I)$, $s = 100$

	$N = 2$	$N = 4$	$N = 6$	$N = 10$	$N = 20$	$N = 50$
$\langle \lambda_{max} \rangle$	0.799 ± 0.806	2.32 ± 1.03	2.46 ± 0.66	3.05 ± 0.56	3.35 ± 0.35	3.70 ± 0.23
$4 - \langle \lambda_{max} \rangle$	3.20	1.68	1.54	0.946	0.655	0.298

Empirical law: $4 - \langle \lambda_{max} \rangle \approx 5.08 N^{-0.713}$

Table 16: The uniform norm of the difference between $\widehat{F}(u)$ and $F(u)$: $N(0, I)$, $y = 1$, $s = 100$

	$\widehat{F} - F\| : 1$	$\|\widehat{F} - F\| : 2$	$\langle \|\widehat{F} - F\| \rangle$
$N = 2$	0.607	0.500	0.596 ± 0.137
$N = 4$	0.438	0.250	0.356 ± 0.081
$N = 6$	0.254	0.223	0.251 ± 0.057
$N = 10$	0.213	0.218	0.174 ± 0.041
$N = 20$	0.0862	0.0640	0.0970 ± 0.0169
$N = 50$	0.0562	0.0539	0.0458 ± 0.0076

Empirical law: $\langle \|\widehat{F} - F\| \rangle \approx 1.060 N^{-0.799}$

Table 17: Local difference between $\widehat{F}(1)$ and $F(1)$: $N(0, I)$, y=1, $s = 100$

	$N = 4$	$N = 6$	$N = 10$	$N = 20$	$N = 50$
$\langle \widehat{F}(1) \rangle$	0.683 ± 0.169	0.680 ± 0.105	0.643 ± 0.062	0.632 ± 0.034	0.616 ± 0.015
$F(1)$	0.609	0.609	0.609	0.609	0.609
$\langle \widehat{F}(1) \rangle - F(1)$	0.0735	0.0710	0.0340	0.0230	0.00680

Empirical law: $\langle \widehat{F}(1) \rangle - F(t) \approx 0.370 N^{-0.996}$

Table 18: Girko's G-estimator of the function $\eta(t)$: $N(0, I)$, $s = 100$

	$y = 1$ $t = 1$	$y = 1$ $t = 2$	$y = 1/2$ $t = 1$	$y = 1/2$ $t = 2$
$N = 2$	0.786 ± 0.114	0.683 ± 0.131	0.738 ± 0.265	0.571 ± 0.309
$N = 4$	0.630 ± 0.087	0.537 ± 0.095	0.608 ± 0.129	0.462 ± 0.136
$N = 6$	0.612 ± 0.053	0.470 ± 0.068	0.571 ± 0.095	0.430 ± 0.099
$N = 10$	0.564 ± 0.039	0.412 ± 0.051	0.541 ± 0.051	0.386 ± 0.071
$N = 20$	0.529 ± 0.022	0.379 ± 0.025	0.520 ± 0.026	0.359 ± 0.027
$N = 50$	0.514 ± 0.009	0.351 ± 0.011	0.508 ± 0.012	0.342 ± 0.012
$\eta(t)$	0.500	0.333	0.500	0.333
$\langle\widehat{\eta}(t)\rangle - \eta(t) \approx$	$0.54N^{-0.94}$	$0.70N^{-0.93}$	$0.47N^{-1.05}$	$0.54N^{-1.02}$