Linear statistics for sample covariance matrices without independent structure in columns

Anna Lytova

University of Alberta, Canada

May 12, 2016

Sample Covariance Matrices

- Basic problem in *multivariate statistics*: by sampling from a high-dimensional distribution, determine its covariance matrix Σ. This problem arises in signal processing, financial mathematics, pattern recognition, computational convex geometry.

- Let $\mathbf{y} = (y_1, \cdots, y_n)$ be a random vector distributed according to μ, $\mathbf{E}\mathbf{y} = 0$. The covariance matrix of $\mathbf{y}$ is defined as

$$\Sigma = \mathbf{E}\mathbf{y}\mathbf{y}^T = \left\{\mathbf{E} y_j y_k\right\}_{j,k=1}^{n} = \left\{\mathbf{Cov}\, y_j y_k\right\}_{j,k=1}^{n}.$$

- We take m independent points $\mathbf{y}_1, \ldots, \mathbf{y}_m$ from the distribution μ and form the sample covariance matrix

$$\Sigma_n = \frac{1}{m} \sum_{\alpha=1}^{m} \mathbf{y}_\alpha \mathbf{y}_\alpha^T.$$

Let $\mathbf{y}_1,\ldots,\mathbf{y}_m$ be m independent zero mean random vectors,

$$\mathbf{y}_1 = \begin{pmatrix} y_{11} \\ \vdots \\ y_{n1} \end{pmatrix}, \quad \ldots, \mathbf{y}_m = \begin{pmatrix} y_{1m} \\ \vdots \\ y_{nm} \end{pmatrix}$$

$$M_n = \sum_{\alpha=1}^{m} \mathbf{y}_\alpha \mathbf{y}_\alpha^T = YY^T, \quad Y = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \ldots & \mathbf{y}_m \end{bmatrix}.$$

The case when all entries are independent is very well studied.

The case of vectors with dependent components:

- Marchenko, Pastur (1967): vectors uniformly distributed on unit sphere.
- Silverstein, Bai (1995): considered the case $\mathbf{y} = \Sigma^{1/2}\mathbf{x}$, where $\mathbf{x}$ consists of iid entries.
- Asymptotic convex geometry: vectors uniformly distributed on convex bodies (results quantitative estimates for extreme eigenvalues)

$$M_n = \sum_{\alpha=1}^{m} \tau_\alpha \mathbf{y}_\alpha \mathbf{y}_\alpha^T :$$

- $m/n \to c \in (0,\infty), \quad n \to \infty$
- $\sigma_m(\Delta) = \sharp\{\alpha : \tau_\alpha \in \Delta\}/m$: $\quad \sigma_m \to \sigma$ weakly as $n \to \infty$
- $\mathbf{y}_1, \ldots, \mathbf{y}_m$ are i.i.d. copies of a normalized isotropic vector $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$, satisfying $\mathbf{E} y_j = 0$ and $\mathbf{E}(y_j y_k) = \delta_{jk}/n$.

Questions:

- Convergence of the normalized counting measures of eigenvalues.
- CLT for linear eigenvalue statistics.
- The same questions for a tensor-product version of sample covariance matrices.

Some Notations

- **Normalized Counting Measure of eigenvalues**:

$$N_n(\Delta) = \sharp\{i : \lambda_i^{(n)} \in \Delta\}/n$$

- **Linear Eigenvalue Statistic** for a given test-function $\varphi : \mathbb{R} \to \mathbb{C}$:

$$\mathcal{N}_n[\varphi] := \sum_{j=1}^{n} \varphi(\lambda_j^{(n)}) = \mathrm{Tr}\varphi(M_n) = \int \varphi(\lambda) n dN_n(\lambda)$$

- **Stieltjes transform** of N_n:

$$s_n(z) = \int \frac{N_n(d\lambda)}{\lambda - z}, \quad \mathrm{Im}\, z \neq 0$$

Note that

$$s_n(z) = n^{-1}\mathrm{Tr}\, G(z),$$

where $G(z) = (M_n - zI)^{-1}$ is the resolvent of M_n.

$$M_n = \sum_{\alpha=1}^{m} \tau_\alpha \mathbf{y}_\alpha \mathbf{y}_\alpha^T :$$

- $m/n \to c \in (0, \infty), \quad n \to \infty$
- $\sigma_m(\Delta) = \sharp\{\tau_\alpha \in \Delta\}/m$: $\quad \sigma_m \to \sigma$ weakly as $n \to \infty$
- $\mathbf{y}_1,\ldots, \mathbf{y}_m$ are i.i.d. copies of a normalized isotropic vector $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$, satisfying $\mathbf{E}y_j = 0$ and $\mathbf{E}(y_j y_k) = \delta_{jk}/n$.

Theorem 1 (Marchenko, Pastur (1967)). If $\mathbf{y}_1,\ldots,\mathbf{y}_m$ are iid random vectors uniformly distributed on the unit sphere in $\mathbb{R}^n$, then N_n converge weakly in probability to non-random N_{MP}. The Stieltjes transform s of N_{MP} is uniquely determined by the functional equation

$$zs(z) = c - 1 - c \int (1 + \tau s(z))^{-1} \sigma(d\tau)$$

considered in the class of functions analytic in $\mathbb{C}\backslash\mathbb{R}$ and such that $\mathrm{Im}\, z \,\mathrm{Im}\, f(z) \geq 0$, $\mathrm{Im}\, z \neq 0$.

Götze, Tikhomirov (2004), Hachem, Loubaton, Najim (2005), Anderson, Zeitouni (2008), Bai, Zhou (2008), Pajor, Pastur (2008), Adamczak (2011), Merlevede, Peligrad (2014), Yaskov (2015)

$$M_n = \sum_{\alpha=1}^{m} \tau_\alpha \mathbf{y}_\alpha \mathbf{y}_\alpha^T :$$

Theorem (Pajor, Pastur (2008)). Let $\mathbf{y}_1,\dots,\mathbf{y}_m$ be i.i.d. copies of a normalized isotropic vector $\mathbf{y}$ satisfying

$$\mathbf{Var}(A_n\mathbf{y},\mathbf{y}) \le ||A_n||^2\delta_n, \quad \delta_n = o(1),\ n \to \infty \tag{1}$$

for any $n \times n$ complex matrix A_n, which does not depend on $\mathbf{y}$. Then NCMs converge weakly in probability to N_{MP} defined in Theorem 1.

Remark. If $\tau_\alpha > 0$, condition (1) can be replaced with

$$\mathbf{E}|(A_n\mathbf{y},\mathbf{y}) - n^{-1}\operatorname{Tr} A_n| \le ||A_n||\delta_n, \quad \delta_n = o(1),\ n \to \infty.$$

Bai, Zhou: If $\mathbf{E}y_j y_k = \Sigma_{jk}$, then the sufficient condition of convergence to N_{MP} is

$$\big((A_n\mathbf{y},\mathbf{y}) - \operatorname{Tr}\Sigma A_n\big)/n \underset{n\to\infty}{\to} 0 \quad \text{in probability.}$$

Main steps of the proof.

$$M_n = \sum_{\alpha=1}^{m} \tau_\alpha \mathbf{y}_\alpha \mathbf{y}_\alpha^T, \quad G(z) = (M_n - zI)^{-1}, \quad s_n = n^{-1}\mathrm{Tr}\, G,$$
$$M_n^\alpha = M_n|_{\tau_\alpha=0}, \quad G^\alpha(z) = (M_n^\alpha - zI)^{-1}, \quad s_n^\alpha = n^{-1}\mathrm{Tr}\, G^\alpha.$$

$$s_n \longrightarrow s?$$

- rank-one perturbation formula

$$G - G^\alpha = -\frac{\tau_\alpha G^\alpha \mathbf{y}_\alpha \mathbf{y}_\alpha^T G^\alpha}{1 + \tau_\alpha (G^\alpha \mathbf{y}_\alpha, \mathbf{y}_\alpha)}$$

- mean and variance of a bilinear form

$$\mathbf{E}(G^\alpha \mathbf{y}_\alpha, \mathbf{y}_\alpha) = \mathbf{E}s_n + O(n^{-1}),$$
$$\mathbf{Var}(G^\alpha \mathbf{y}_\alpha, \mathbf{y}_\alpha) = o(1), \; n \to \infty.$$

$$zG = -1 + GM,$$

$$z\mathbf{E}s_n = -1 + \frac{m}{n} - \frac{1}{n}\sum_{\alpha=1}^{m} \mathbf{E}\frac{1}{1+\tau_\alpha(G^\alpha \mathbf{y}_\alpha, \mathbf{y}_\alpha)}$$

$$= -1 + \frac{m}{n} - \frac{1}{n}\sum_{\alpha=1}^{m} \frac{1 + r_{\alpha n}}{1+\tau_\alpha \mathbf{E}s_n^\alpha(z)},$$

where $r_{\alpha n} = \tau_\alpha \mathbf{E}\{(G^\alpha \mathbf{y}_\alpha, \mathbf{y}_\alpha)^\circ/(1+\tau_\alpha(G^\alpha \mathbf{y}_\alpha, \mathbf{y}_\alpha))\} = o(1)$. Hence

$$zs(z) = -1 + c - c\int \frac{\sigma(d\tau)}{1+\tau s(z)}. \quad \square$$

Remark

$$\frac{1}{A} = \frac{1}{\mathbf{E}A} - \frac{A^\circ}{A\mathbf{E}A}, \quad A^\circ = A - \mathbf{E}A$$

$$\frac{1}{A} = \frac{1}{\mathbf{E}A} - \frac{A^\circ}{(\mathbf{E}A)^2} + \frac{(A^\circ)^2}{(\mathbf{E}A)^3} - \ldots + (-1)^t \frac{(A^\circ)^t}{A(\mathbf{E}A)^t}.$$

Theorem (Pajor, Pastur (2008)). Let $\mathbf{y}_1,...,\mathbf{y}_m$ be i.i.d. copies of a random vector $\mathbf{y}$ satisfying

$$\mathbf{Var}(A_n\mathbf{y},\mathbf{y}) \le ||A_n||^2\delta_n, \quad \delta_n = o(1),\ n \to \infty$$

for any $n \times n$ complex matrix A_n, which does not depend on $\mathbf{y}$. Then NCMs converge weakly in probability to N_{MP} defined in Theorem 1.

- Normalized isotropic vectors with a log-concave distribution satisfy the conditions of the theorem.
- It follows from the theorem that if

$$n^{-1}\mathcal{N}_n[\varphi] = \int \varphi(\lambda)N_n(d\lambda) = n^{-1}\mathrm{Tr}\,\varphi(M_n)$$

is a normalized linear eigenvalue statistic corresponding to a bounded continuous test-function $\varphi : \mathbb{R} \to \mathbb{C}$, then in probability:

$$\lim_{m,n\to\infty,m/n\to c} n^{-1}\mathcal{N}_n[\varphi] = \int \varphi(\lambda)N_{MP}(d\lambda).$$

$$\mathbf{y}_\alpha - ?:\ \mathcal{N}_n^\circ[\varphi] \underset{m,n\to\infty,m/n\to c}{\longrightarrow} \mathcal{N}(0,V) \textit{ in distribution}$$

CLT for linear eigenvalue statistics

Girko (2001), Bai, Silverstein (2004), Chatterjee, Bose (2004), Anderson, Zeitouni (2006), Pan, Zhou (2008), AL, Pastur (2009), Shcherbina (2011), Najim, Yao (2013)

$$M_n = M_n(\mathbf{y}) = \sum_{\alpha=1}^{m} \tau_\alpha \mathbf{y}_\alpha \mathbf{y}_\alpha^T$$

The CLT for linear eigenvalue statistics for $M_n(\mathbf{y})$ is expected to be valid, in particular, if $\mathbf{y}$ is a normalized isotropic vector with a log-concave distribution.

Definition. (i) A measure μ is *log-concave* if

$$\mu(\theta A + (1-\theta)B) \geq \mu(A)^\theta \mu(B)^{(1-\theta)}, \quad \theta \in [0,1].$$

(ii) The distribution of random vector $\mathbf{y} \in \mathbb{R}^n$ is called *unconditional* if its components $\{y_j\}_{j=1}^n$ have the same joint distribution as $\{\pm y_j\}_{j=1}^n$ for any choice of signs.

Guédon, AL, Pajor, Pastur (2013):

Lemma. If $\mathbf{y}$ has an unconditional distribution,

$$a_{2,2} := \mathbf{E} y_i^2 y_j^2 = n^{-2} + O(n^{-3}),\ i \neq j,$$
$$\kappa_4 := \mathbf{E} y_j^4 - 3a_{2,2} = O(n^{-2}),$$

and $\varphi \in H_{2+\delta}$, where $H_s = \{\psi : ||\psi||_s^2 = \int (1+|\xi|)^{2s} |\widehat{\psi}(\xi)|^2 d\xi < \infty\}$, then

$$\mathbf{Var} \mathcal{N}_n[\varphi] \leq C ||\varphi||_{2+\delta}^2,$$

where C is an absolute constant.

- Shcherbina (2011):

$$\mathbf{Var}\{\mathcal{N}_n[\varphi]\} \leq C_q ||\varphi||_q^2 \int_0^\infty d\eta e^{-\eta} \eta^{2q-1} \int_{-\infty}^\infty \mathbf{Var}\{\text{Tr } G(x+i\eta)\} dx.$$

- Normalized isotropic vectors with an unconditional log-concave distribution satisfy the conditions of the lemma.

Definition. We say that a normalized isotropic vector $\mathbf{y} \in \mathbb{R}^n$ is a CLT-vector if it has an unconditional distribution and satisfies the following moment conditions:

- there exist n-independent $a, b \in \mathbb{R}$, s.t.

$$a_{2,2} := \mathbf{E} y_i^2 y_j^2 = n^{-2} + a n^{-3} + o(n^{-3}), \quad i \neq j, n \to \infty,$$
$$\kappa_4 := \mathbf{E} y_j^4 - 3a_{2,2} = b n^{-2} + o(n^{-2}), \quad n \to \infty,$$

- for any $n \times n$ complex matrix A_n, which does not depend on $\mathbf{y}$,

$$\mathbf{E}|(A_n \mathbf{y}, \mathbf{y})^\circ|^4 \leq ||A_n||^4 \widetilde{\delta}_n, \quad \widetilde{\delta}_n = O(n^{-2}), \; n \to \infty$$

Example. If $\mathbf{x}$ is uniformly distributed on $B_p^n = \{\mathbf{x} \in \mathbb{R}^n : \; ||\mathbf{x}||_p = (\sum |x_j|^p)^{1/p} \leq 1\}$, then

$$\mathbf{y} = \left(\frac{1}{n} \frac{B(1/p, 2/p)}{B(n/p + 1, 2/p)} \right)^{1/2} \mathbf{x}$$

is a CLT-vector, and

$$a = -4p^{-1}, \quad b = \Gamma(1/p)\Gamma(5/p)\Gamma(3/p)^{-2} - 3.$$

Theorem (Guédon, AL, Pajor, Pastur (2013)).
Let $M_n = \sum_{\alpha=1}^{m} \tau_\alpha \mathbf{y}_\alpha \mathbf{y}_\alpha{}^T$, where

- $\{\mathbf{y}_\alpha\}_{\alpha=1}^m$ are i.i.d. CLT-vectors,
- $\int \tau^4 d\sigma(\tau) < \infty$.

If $\varphi \in H_{2+\varepsilon}$, $\varepsilon > 0$, then $\mathcal{N}_n^\circ[\varphi] = \mathcal{N}_n[\varphi] - \mathbf{E}\{\mathcal{N}_n[\varphi]\}$ converges in distribution to a Gaussian random variable with zero mean and the variance

$$V[\varphi] = \lim_{\eta \downarrow 0} \frac{1}{2\pi^2} \int \varphi(\mu) d\mu \int \varphi(\lambda) d\lambda \, \mathrm{Re} \left[C(z_\mu, \overline{z_\lambda}) - C(z_\mu, z_\lambda) \right],$$

where $z_\lambda = \lambda + \eta$, $z_\mu = \mu + \eta$,

$$\begin{aligned} C(z_1, z_2) &= \lim_{n\to\infty} \mathbf{Cov}\{\mathrm{Tr} G(z_1),\ \mathrm{Tr} G(z_2)\} \\ &= \frac{\partial^2}{\partial z_1 \partial z_2} \left(-(a+b) s(z_1) s(z_2) \frac{\Delta z}{\Delta s} + 2 \ln \frac{\Delta s}{\Delta z} \right), \end{aligned}$$

$$zs(z) = c - 1 - c \int (1 + \tau s(z))^{-1} \sigma(d\tau), \quad \Delta s = s(z_1) - s(z_2),\ \Delta z = z_1 - z_2.$$

If $\tau_1 = ... = \tau_m = 1$, then

$$V[\varphi] = \frac{1}{2\pi^2} \int_{a_-}^{a_+} \int_{a_-}^{a_+} \left(\frac{\triangle \varphi}{\triangle \lambda} \right)^2 \frac{(4c - (\lambda_1 - a_m)(\lambda_2 - a_m)) d\lambda_1 d\lambda_2}{\sqrt{(a_+ - \lambda_1)(\lambda_1 - a_-)}\sqrt{(a_+ - \lambda_2)(\lambda_2 - a_-)}} + \frac{a+b}{4c\pi^2} \left(\int_{a_-}^{a_+} \varphi(\mu) \frac{\mu - a_m}{\sqrt{(a_+ - \mu)(\mu - a_-)}} d\mu \right)^2, \quad (2)$$

where $a_\pm = (1 \pm \sqrt{c})^2$, $\quad a_m = 1 + c$, and as before a, b:

$$a_{2,2} = \mathbf{E} y_{\alpha i}^2 y_{\alpha j}^2 = n^{-2} + a n^{-3} + o(n^{-3}),$$
$$\mathbf{E} y_{\alpha j}^4 - 3 a_{2,2} = b n^{-2} + o(n^{-2}).$$

The expression for $V[\varphi]$ coincides with the known one for for i.i.d. case, in which $a = 0$. In particular, we can get (2) if we suppose that all $\{y_{\alpha j}\}_{\alpha,j=1}^{m,n}$ are i.i.d. random variables having finite forth moment and

$$\mathbf{E} y_{\alpha j} = 0, \quad \mathbf{E} y_{\alpha j}^2 = n^{-1} + a n^{-2}/2 + o(n^{-3}), \quad b = \mathbf{E} y_{\alpha j}^4 - 3(\mathbf{E} y_{\alpha j}^2)^2.$$

Main steps of the proof. Let $Z_n[\varphi](x) = \mathbf{E}\exp(ix\mathcal{N}_n^\circ[\varphi])$.
We show that

$$\lim_{n\to\infty} Z_n = Z, \quad \lim_{n\to\infty} Z_n' = -xV[\varphi]Z.$$

- **Approximation procedure.** For $\varphi \in H_{2+\delta}$, $\eta > 0$, put

$$\varphi_\eta = \varphi * P_\eta, \quad P_\eta = \frac{1}{\pi}\frac{\eta}{x^2+\eta^2}.$$

  Then

$$|Z_n[\varphi] - Z_n[\varphi_\eta]| \le |x|(\mathbf{Var}\mathcal{N}_n[\varphi-\varphi_\eta])^{1/2} \le |x|||\varphi-\varphi_\eta||_{2+\delta} \underset{\eta\to 0}{\to} 0.$$

  Hence uniformly in $|x| \le C$: $\lim_{n\to\infty} Z_n[\varphi] = \lim_{\eta\downarrow 0}\lim_{n\to\infty} Z_n[\varphi_\eta]$.
- $$\mathcal{N}_n^\circ[\varphi_\eta] = \frac{1}{\pi}\int \varphi(\mu)\,\mathrm{Im}\,\mathrm{Tr}\,G(\mu+i\eta)^\circ d\mu,$$

$$\mathrm{Tr}\,G(\mu+i\eta)^\circ = -\frac{1}{z}\sum_{\alpha=1}^{m}\Big(\frac{1}{1+\tau_\alpha(G^\alpha\mathbf{y}_\alpha,\mathbf{y}_\alpha)}\Big)^\circ$$

$$\frac{1}{A} = \frac{1}{\mathbf{E}A} - \frac{A^\circ}{(\mathbf{E}A)^2} + \frac{(A^\circ)^2}{(\mathbf{E}A)^3} - \frac{(A^\circ)^3}{A(\mathbf{E}A)^3}.$$

Tensor Product Version of Sample Covariance Matrices

Consider random vectors of the form

$$\mathbf{Y} = \mathbf{y}^{(1)} \otimes \ldots \otimes \mathbf{y}^{(k)} \in (\mathbb{R}^n)^{\otimes k},$$

where $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(k)}$ are i.i.d. copies of a random vector $\mathbf{y} \in \mathbb{R}^n$,

$$Y_{\mathbf{j}} = y^{(1)}_{j_1} \times \ldots \times y^{(k)}_{j_k}, \quad \mathbf{j} = \{j_1, \ldots, j_k\},\ 1 \le j_\ell \le n,$$

k is fixed. Let $\mathcal{M}_n$ be an $n^k \times n^k$ random matrix of the form

$$\mathcal{M}_n = \mathcal{M}_{m,n,k}(\mathbf{y}) = \sum_{\alpha=1}^{m} \tau_\alpha \mathbf{Y}_\alpha \mathbf{Y}_\alpha{}^T,$$

- $m/n^k \to c \in (0, \infty), \quad n \to \infty$
- $\sigma_m(\Delta) = \sharp\{\tau_\alpha \in \Delta\}/m$: $\quad \sigma_m \to \sigma$ weakly as $n \to \infty$
- $\mathbf{Y}_\alpha = \mathbf{y}^{(1)}_\alpha \otimes \ldots \otimes \mathbf{y}^{(k)}_\alpha$, and $\{\mathbf{y}^{(p)}_\alpha\}_{\alpha,p=1}^{m,k}$ are i.i.d. random vectors in $\mathbb{R}^n$

$$\mathcal{M}_n(\mathbf{y}) = \sum_{\alpha=1}^{m} \tau_\alpha \mathbf{Y}_\alpha {\mathbf{Y}_\alpha}^T, \quad \mathbf{Y}_\alpha = \mathbf{y}_\alpha^{(1)} \otimes ... \otimes \mathbf{y}_\alpha^{(k)}$$

was studied by M. Hastings et al as a model of data hiding and correlation scheme of Quantum Informatics (Ambainis, A., Harrow, A. W., and Hastings, M. B. (2012). *Random tensor theory: extending random matrix theory to random product states.* Commun. Math. Phys., **310** 1, 25-74).

Theorem. If $\mathbf{y}$ satisfies

$$\mathbf{Var}|(A_n\mathbf{y}, \mathbf{y})^\circ| \leq ||A_n||^2 \delta_n, \quad \delta_n = O(n^{-\alpha}),\ n \to \infty.$$

for any $n \times n$ complex matrix A_n, which does not depend on $\mathbf{y}$, then the NCMs of $\mathcal{M}_n(\mathbf{y})$ converges weakly in probability to N_{MP}.

Lemma. If $\mathbf{E}|(A_n\mathbf{y}, \mathbf{y})^\circ|^t \leq C_t ||\mathcal{A}_n||^t \delta_n$, then

$$\mathbf{E}|(\mathcal{A}_n\mathbf{Y}, \mathbf{Y})^\circ|^{2t} \leq C_t k^{t/2} ||\mathcal{A}_n||^t \delta_n$$

for any $n^k \times n^K$ complex matrix $\mathcal{A}_n$, which does not depend on $\mathbf{y}$.

$$\mathcal{M}_n(\mathbf{y}) = \sum_{\alpha=1}^{m} \tau_\alpha \mathbf{Y}_\alpha {\mathbf{Y}_\alpha}^T, \quad \mathbf{Y}_\alpha = \mathbf{y}_\alpha^{(1)} \otimes \ldots \otimes \mathbf{y}_\alpha^{(k)}$$

Lemma. If $\varphi \in H_s$, $s > 5/2$, and $\mathbf{Var}(A_n\mathbf{y}, \mathbf{y}) \le C||A_n||^2 n^{-1}$, then

$$\mathbf{Var}\{\mathcal{N}_n[\varphi]\} = O(kn^{k-1}), \; n \to \infty.$$

Definition. **($k = 2$).** We say that a normalized isotropic vector $\mathbf{y} \in \mathbb{R}^n$, is a *CLT-vector* if it has unconditional permutationally invariant distributions and satisfy the following conditions:

$$\begin{aligned}
a_{2,2} &= \mathbf{E} y_i^2 y_j^2 = n^{-2} + an^{-3} + o(n^{-3}), \quad i \neq j, n \to \infty, \\
\kappa_4 &= \mathbf{E} y_j^4 - 3a_{2,2} = bn^{-2} + o(n^{-2}), \quad n \to \infty, \\
a_{2,2,2} &:= \mathbf{E} y_i^2 y_j^2 y_k^2 = n^{-3} + O(n^{-4}), \\
a_{2,4} &:= \mathbf{E} y_i^2 y_j^4 = O(n^{-3}), \quad a_6 := \mathbf{E} y_i^6 = O(n^{-3}),
\end{aligned}$$

and for every $n \times n$ matrix A_n which does not depend on $\mathbf{y}$,

$$\mathbf{E}|(A_n\mathbf{y}, \mathbf{y})^\circ|^6 \le C||A_n||^6 n^{-3}.$$

Theorem (CLT). If $\mathcal{M}_n = \sum_{\alpha=1}^{m} \tau_\alpha \mathbf{Y}_\alpha {\mathbf{Y}_\alpha}^T$, $\mathbf{Y}_\alpha = \mathbf{y}_\alpha^{(1)} \otimes \mathbf{y}_\alpha^{(2)}$, and $\{\mathbf{y}_\alpha^{(p)}\}_{\alpha,p=1}^{m,2}$ are i.i.d. copies of a CLT-vector $\mathbf{y}$, and $\varphi \in H_s$, $s > 5/2$, then $n^{-1/2}\mathcal{N}_n^\circ[\varphi]$ converges in distribution to a Gaussian random variable with zero mean and the variance

$$V[\varphi] = \lim_{\eta \downarrow 0} \frac{1}{2\pi^2} \int \varphi(\mu) d\mu \int \varphi(\lambda) d\lambda \, \mathrm{Re}\left[C(z_\mu, \overline{z_\lambda}) - C(z_\mu, z_\lambda)\right],$$

where $z_\lambda = \lambda + \eta$, $z_\mu = \mu + \eta$,

$$C(z_1, z_2) = -k(a+b+2) \frac{\partial^2}{\partial z_1 \partial z_2} \left(s(z_1) s(z_2) \frac{\Delta z}{\Delta s} \right),$$

$$zs(z) = c - 1 - c \int (1 + \tau s(z))^{-1} \sigma(d\tau).$$

Remark. If $\{\mathbf{y}_\alpha^{(p)}\}_{\alpha,p=1}^{m,k}$ are i.i.d. copies of random vector $\mathbf{y}$ uniformly distributed on the unit sphere in $\mathbb{R}^n$, then

$$a = -2, \quad b = 0.$$

After renormalization, the CLT is expected to be valid for $\mathcal{N}_n^\circ[\varphi]$.

Thank you!