Extended Avanti's CLT Theorem

Runze

July 27, 2015

All notations in this writeup are according to the latest version of ***A limit theorem for scaled eigenvectors of random dot product graphs***.

Corollary 0.1 *Let $X_1, \cdots, X_n$ be mean-zero independent random matrices, defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with values in $\mathbb{C}^{d\times d}_{\text{Herm}}$ and such that there exists a $M > 0$ with $\|X_i\| \leq M$ almost surely for all $1 \leq i \leq m$. Define $\sigma^2 \equiv \lambda_{\max}\left(\sum_{i=1}^n \mathbb{E}[X_i^2]\right)$. Then for all $t \geq 0$,*

$$\mathbb{P}\left(\lambda_{\max}\left(\sum_{i=1}^n X_i\right) \geq t\right) \leq de^{-\frac{t^2}{8\sigma^2+4Mt}},$$

and

$$\mathbb{P}\left(\left\|\sum_{i=1}^n X_i\right\| \geq t\right) \leq 2de^{-\frac{t^2}{8\sigma^2+4Mt}}.$$

Remark: See Corollary 7.1 in http://arxiv.org/pdf/0911.0600v2.pdf.

Theorem 0.2 *Let $P : [n]^2 \to [0, 1]$ be symmetric, and $A^{(1)}, \cdots, A^{(m)} \overset{iid}{\sim} \text{Bern}(P)$ also be symmetric. Define $\bar{A} = \frac{1}{m}\sum_{t=1}^m A^{(t)}$, then for any constant $c > 0$ there exists another constant $n_0(c)$, independent of n or P, such that if $n > n_0$, then for all η satisfying $n^{-c} \leq \eta \leq 1/2$,*

$$\mathbb{P}\left(\|\bar{A} - P\| \leq 4\sqrt{n\ln(n/\eta)}\right) \geq 1 - \eta.$$

Remark: This is the extended version of Theorem 3.1 in http://arxiv.org/pdf/0911.0600v2.pdf.
Proof: Let $\{e_i\}_{i=1}^n$ be the canonical basis for $\mathbb{R}^n$. For each $1 \leq i, j \leq n$, define a corresponding matrix G_{ij}:

$$G_{ij} \equiv \begin{cases} e_i e_j^T + e_j e_i^T, & i \neq j; \\ e_i e_i^T, & i = j. \end{cases}$$

Thus $A^{(t)} = \sum_{1\leq i\leq j\leq n} A_{ij}^{(t)} G_{ij}$ for $1 \leq t \leq m$, $\bar{A} = \sum_{1\leq i\leq j\leq n} \frac{1}{m}\sum_{t=1}^m A_{ij}^{(t)} G_{ij}$ and $P = \sum_{1\leq i\leq j\leq n} P_{ij} G_{ij}$. Then we have $\bar{A} - P = \sum_{1\leq i\leq j\leq n} X_{ij}$, where $X_{ij} \equiv \left(\frac{1}{m}\sum_{t=1}^m A_{ij}^{(t)} - P_{ij}\right) G_{ij}$, $1 \leq i \leq j \leq n$.

Notice that random matrices X_{ij} are independent and have mean zero. Also,

$$\|X_{ij}\| \leq \|G_{ij}\| = 1$$

as the eigenvalues of G_{ij} are always contained in the set $\{-1, 0, 1\}$. Thus the assumptions of Corollary 0.1 apply with $M = 1$.

Since

$$\begin{aligned}
\mathbb{E}[X_{ij}^2] =& Var(X_{ij}) \\
=& Var\left(\frac{1}{m}\sum_{t=1}^{m} A_{ij}^{(t)} - P_{ij}\right) G_{ij}^2 \\
=& Var\left(\frac{1}{m}\sum_{t=1}^{m} A_{ij}^{(t)}\right) G_{ij}^2 \\
=& \frac{1}{m} Var(A_{ij}^{(t)}) G_{ij}^2 \\
=& \frac{1}{m} P_{ij}(1 - P_{ij}) G_{ij}^2
\end{aligned}$$

and

$$G_{ij}^2 \equiv \begin{cases} e_i e_i^T + e_j e_j^T, & i \neq j; \\ e_i e_i^T, & i = j. \end{cases}$$

Therefore,

$$\begin{aligned}
\sum_{i \leq j} \mathbb{E}[X_{ij}^2] &= \sum_{i} \frac{1}{m} P_{ii}(1 - P_{ii}) e_i e_i^T + \sum_{i<j} \frac{1}{m} P_{ij}(1 - P_{ij})(e_i e_i^T + e_j e_j^T) \\
&= \sum_{i=1}^{n} \left(\frac{1}{m}\sum_{j=1}^{n} P_{ij}(1 - P_{ij})\right) e_i e_i^T.
\end{aligned}$$

This is a diagonal matrix. Define $\Delta \equiv \max_{i\in[n]} d_P(i)$, then its largest eigenvalue is at most

$$\max_{i\in[n]} \left(\frac{1}{m}\sum_{j=1}^{n} P_{ij}(1 - P_{ij})\right) \leq \max_{i\in[n]} \frac{1}{m}\sum_{j=1}^{n} P_{ij} = \Delta/m.$$

Apply Corollary 0.1 with $\sigma^2 = \Delta/m$ and $M = 1$, we get for any $t > 0$, $m \geq 1$,

$$\mathbb{P}(\|\bar{A} - P\| \geq t) \leq 2ne^{-\frac{t^2}{8\Delta/m+4t}} \leq 2ne^{-\frac{t^2}{8n+4t}}. \tag{1}$$

Now let $c > 0$ be given and assume $n^{-c} \leq \eta \leq 1/2$. Then there exists a $n_0(c)$ independent of n and P such that whenever $n > n_0(c)$,

$$t = 4\sqrt{n \ln(2n/\eta)} \leq 2n.$$

Plugging this t into Equation (1), we get

$$\mathbb{P}(\|\bar{A} - P\| \geq 4\sqrt{n \ln(2n/\eta)}) \leq 2ne^{-\frac{t^2}{16n}} = 2ne^{-\frac{16n\ln(2n/\eta)}{16n}} = \eta.$$

■

Proposition 0.3 *It holds with probability greater than* $1 - \frac{2}{n^2}$ *that*

$$\|\bar{A}^2 - P^2\|_F \leq \sqrt{2n^3 \log n}.$$

Remark: This is Proposition 4.2 in http://arxiv.org/pdf/1207.6745.pdf.
Proof: Notice that $\bar{A}^2_{ij} = \sum_{k=1}^n \bar{A}_{ik}\bar{A}_{kj}$ is a sum of n independent random variables between 0 and 1. Since $E[\bar{A}^2_{ij}] = P^2_{ij}$, by Hoeffding's inequality, we have

$$\mathbb{P}\left(|A^2_{ij} - P^2_{ij}| > \sqrt{2n\log n}\right) \leq 2\exp\left(-\frac{4n\log n}{n}\right) = \frac{2}{n^4}.$$

Using a union bound, we have

$$\begin{aligned}\mathbb{P}\left(\|A^2 - P^2\|_F > \sqrt{2n^3\log n}\right) &\leq \mathbb{P}\left(\cup_{i,j}\left\{|A^2_{ij} - P^2_{ij}| > \sqrt{2n\log n}\right\}\right)\\ &\leq \sum_{i,j}\mathbb{P}\left(|A^2_{ij} - P^2_{ij}| > \sqrt{2n\log n}\right)\\ &= \sum_{i,j}\frac{2}{n^4} = \frac{2}{n^2}.\end{aligned}$$

■

Proposition 0.4 *Assuming that the matrix $\mathbb{E}[XX^T]$ is rank d and has distinct eigenvalues. Also we suppose there exists $\delta > 0$ such that $\delta < \min_{i\neq j}|\lambda_i(\mathbb{E}[XX^T]) - \lambda_j(\mathbb{E}[XX^T])|$ and $\delta < \lambda_d(\mathbb{E}[XX^T]$. For $i, j \leq d+1$, $i \neq j$ and n sufficiently large, it holds with probability greater than $1 - \frac{2d^2}{n^2}$ that*

$$|\lambda_i(P) - \lambda_j(P)| > \delta n/2.$$

Remark: This is Proposition 4.3 in http://arxiv.org/pdf/1207.6745.pdf.

Lemma 0.5 *Let $(X, A^{(1)}, \cdots, A^{(m)}) \sim \mathrm{RDPG}(F)$ and let $\bar{X} = \bar{V}\bar{S}^{1/2}$ be our estimate for X based on $\bar{A} = \frac{1}{m}\sum_{t=1}^m A^{(t)}$. Define $\delta_d = \min\{\min_{1\leq i\leq d-1}|\lambda_i - \lambda_{i+1}|, \lambda_d\}$ where $\lambda_1 > \lambda_2 > \cdots > \lambda_d$ are the ordered, distinct, positive eigenvalues of Δ. For n sufficiently large, following bounds hold with probability greater than $1-\eta$,*

$$\frac{\delta_d n}{2} \leq \|S\| \leq n,$$

$$\frac{\delta_d n}{2} \leq \|\bar{S}\| \leq n.$$

Remark: This is an extended version of Lemma 1 in ***A limit theorem for scaled eigenvectors of random dot product graphs***.
Proof: To be continue. ■

Lemma 0.6 *Let $(X, A^{(1)}, \cdots, A^{(m)}) \sim \mathrm{RDPG}(F)$ and let $\bar{X} = \bar{V}\bar{S}^{1/2}$ be our estimate for X based on $\bar{A} = \frac{1}{m}\sum_{t=1}^m A^{(t)}$. Define $\delta_d = \min\{\min_{1\leq i\leq d-1}|\lambda_i - \lambda_{i+1}|, \lambda_d\}$ where $\lambda_1 > \lambda_2 > \cdots > \lambda_d$ are the ordered, distinct, positive eigenvalues of Δ. Let c be arbitrary. There exists a constant $n_0(c)$ such that if $n > n_0$, then for any η satisfying $n^{-c} < \eta < 1/2$, the following bounds hold with probability greater than $1-\eta$,*

$$\|XX^T - \bar{A}\| \leq 4(n\log(n/\eta))^{1/2}, \tag{2}$$

$$\|\bar{V} - V\|_F \leq 8\delta_d^{-1}n^{-1/2}(2d\log(n/\eta))^{1/2}. \tag{3}$$

$$\|\bar{X} - \tilde{X}\|_F = \|\bar{V}\bar{S}^{1/2} - VS^{1/2}\|_F \leq 16\delta_d^{-3/2}(2d\log(n/\eta))^{1/2}, \tag{4}$$

Remark: This is an extended version of Lemma 1 in ***A limit theorem for scaled eigenvectors of random dot product graphs***.

Proof: By Theorem 0.2, we get Inequality (2).

For Inequality (3), by Davis-Kahan Theorem, Proposition 0.4 and Inequality (2), we have

$$\begin{aligned}\|\bar{V}_i - V_i\|_2 &\leq \frac{\|\bar{A} - P\|_2}{\min\{|\lambda_i(P) - \lambda_{i+1}(P)|, |\lambda_i(P) - \lambda_{i-1}(P)|\}} \\ &\leq \frac{\|\bar{A} - P\|_2}{\delta_d n/2} \\ &\leq 8\delta_d^{-1} n^{-1/2} (\log(n/\eta))^{1/2},\end{aligned}$$

where $\bar{V}_i$ represents the ith column of $\bar{V}$ and V_i represents the ith column of V.

It is not very clear here since Proposition 0.4 happens with probability $1 - 2d^2/n^2$.

Thus

$$\begin{aligned}\|\bar{V} - V\|_F &= \left(\sum_{i=1}^{d} \|\bar{V}_i - V_i\|_2^2\right)^{1/2} \\ &\leq \left(d \cdot \left(8\delta_d^{-1} n^{-1/2} (\log(n/\eta))^{1/2}\right)^2\right)^{1/2} \\ &= 8\delta_d^{-1} n^{-1/2} (d \log(n/\eta))^{1/2}\end{aligned}$$

For Inequality (4), first notice that

$$\begin{aligned}&\|\bar{V}\bar{S}^{1/2} - VS^{1/2}\|_F \\ \leq&\|\bar{V}\bar{S}^{1/2} - \bar{V}S^{1/2}\|_F + \|\bar{V}S^{1/2} - VS^{1/2}\|_F \\ =&\|\bar{V}(\bar{S}^{1/2} - S^{1/2})\|_F + \|(\bar{V} - V)S^{1/2}\|_F.\end{aligned}$$

By Proposition 0.3 and Weyl's inequality, we know

$$\begin{aligned}\lambda_i^2(|\bar{A}|) - \lambda_i^2(P) =&\lambda_i(\bar{A}^2) - \lambda_i(P^2) \\ \leq&\|\bar{A}^2 - P^2\|_2 \\ \leq&\|\bar{A}^2 - P^2\|_F \\ \leq&\sqrt{2n^3 \log n}\end{aligned}$$

with probability greater than $1 - \frac{2}{n^2}$. Thus it holds with probability greater than $1 - \frac{2}{n^2}$ that

$$\begin{aligned}&\lambda_i^{1/2}(|\bar{A}|) - \lambda_i^{1/2}(P) \\ =&\frac{\lambda_i^2(|\bar{A}|) - \lambda_i^2(P)}{\left(\lambda_i(|\bar{A}|) + \lambda_i(P)\right)\left(\lambda_i^{1/2}(|\bar{A}|) + \lambda_i^{1/2}(P)\right)} \\ \leq&\frac{\sqrt{2n^3 \log n}}{(\delta_d n)^{3/2}} = \sqrt{\frac{2 \log n}{\delta_d^3}}.\end{aligned}$$

Combined with Inequality (3) and Lemma 0.5, also notice that $\delta_d \leq 1$, then we have

$$
\begin{aligned}
&\|\bar{V}\bar{S}^{1/2} - VS^{1/2}\|_F \\
\leq&\|\bar{V}(\bar{S}^{1/2} - S^{1/2})\|_F + \|(\bar{V} - V)S^{1/2}\|_F \\
\leq&\|\bar{V}\|_F\|\bar{S}^{1/2} - S^{1/2}\|_2 + \|\bar{V} - V\|_F\|S^{1/2}\|_2 \\
\leq&\sqrt{d} \cdot \sqrt{\frac{2\log n}{\delta_d^3}} + 8\delta_d^{-1}n^{-1/2}(2d\log(n/\eta))^{1/2} \cdot \sqrt{n} \\
=&16\delta_d^{-3/2}(2d\log(n/\eta))^{1/2}.
\end{aligned}
$$

■

Lemma 0.7 *In the setting of Lemma 0.6, with probability greater than* $1 - 2\eta$*,*

$$\|S - \bar{S}\|_F \leq C\delta_d^{-2}d\log(n/\eta).$$

Remark: This is an extended version of Lemma 2 in ***A limit theorem for scaled eigenvectors of random dot product graphs***.
Proof: Replace A with $\bar{A}$. Omitted here. ■

Lemma 0.8 *In the setting of Lemma 0.6, with probability greater than* $1 - 2\eta$*,*

$$\|V^T\bar{V} - I\|_F \leq \frac{Cd\log(n/\eta)}{\delta_d^2 n}.$$

Remark: This is an extended version of Lemma 3 in ***A limit theorem for scaled eigenvectors of random dot product graphs***.
Proof: Replace A with $\bar{A}$. Omitted here. ■

Theorem 0.9 *Let* $(X, A^{(1)}, \cdots, A^{(m)}) \sim \mathrm{RDPG}(F)$ *be d-dimensional random dot product graph, i.e., F is a distribution for points in* $\mathbb{R}^d$*, and let* $\bar{X}$ *be our estimate for X based on* $\bar{A} = \frac{1}{m}\sum_{t=1}^m A^{(t)}$*. We assume F has a diagonal second moment matrix with distinct positive entries along the diagonal. Let* $\Phi(z, \Sigma)$ *denote the cumulative distribution function for the multivariate normal, with mean zero and covariance matrix* Σ*, evaluated at z. Then there exists a sequence of orthogonal matrices* W_n *converging to the identity almost surely such that for each component i and any* $z \in \mathbb{R}^d$*,*

$$\mathbb{P}\{\sqrt{n}(W_n\bar{X}_i - X_i) \leq z\} \to \int \Phi(z, \Sigma(x))dF(x)$$

where $\bar{\Sigma}(x) = \frac{1}{m}\Delta^{-1}E\left[X_jX_j^T(x^TX_j - (x^TX_j)^2)\right)\Delta^{-1}$ *and* $\Delta = E[X_1X_1^T]$ *is the second moment matrix. That is, the sequence of random variables* $\sqrt{n}(W_n\bar{X}_i - X_i)$ *converges in distribution to a mixture of multivariate normals. We denote this mixture by* $\mathcal{N}(0, \Sigma(X_i))$*.*

Remark: This is an extended version of Theorem 1 in ***A limit theorem for scaled eigenvectors of random dot product graphs***.

Proof: Let $\bar{A}VS^{-1/2} = [Y_1, \cdots, Y_n]^T$ where $Y_i \in \mathbb{R}^d$. We first show that $\sqrt{n}(Y_i - \tilde{X}_i)$ is multivariate normal in the limit. We have

$$\begin{aligned}
\sqrt{n}(Y_i - \tilde{X}_i) &= \sqrt{n}\left((\bar{A}\tilde{X}S^{-1})^T_{\cdot i} - (P\tilde{X}S^{-1})^T_{\cdot i}\right) \\
&= nS^{-1}W_n\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n}(\bar{A}_{ij} - P_{ij})X_j\right) \\
&= nS^{-1}W_n\left(\frac{1}{\sqrt{n}}\sum_{j\neq i}(\bar{A}_{ij} - X_i^TX_j)X_j - \frac{X_i^TX_i}{\sqrt{n}}X_i\right).
\end{aligned}$$

Conditioning on $X_i = x_i$, the scaled sum

$$\frac{1}{\sqrt{n}}\sum_{j\neq i}(\bar{A}_{ij} - X_i^TX_j)X_j$$

is a sum of i.i.d, mean zero random variables with covariance matrix

$$\tilde{\Sigma}(x) = \frac{1}{m}E\left[X_jX_j^T(x^TX_j - (x^TX_j)^2)\right).$$

Detailed calculation for the expectation is as following:

$$\begin{aligned}
&E\left[(\bar{A}_{ij} - x_i^TX_j)X_j\right] \\
=&E\left[\frac{1}{m}\sum_{t=1}^{m}A_{ij}^{(t)}X_j - (x_i^TX_j)X_j\right] \\
=&E\left[E\left[\left(\frac{1}{m}\sum_{t=1}^{m}A_{ij}^{(t)}X_j - (x_i^TX_j)X_j\right)|X_j\right]\right] \\
=&E\left[\frac{1}{m}\sum_{t=1}^{m}(x_i^TX_j)X_j - (x_i^TX_j)X_j\right] \\
=&0
\end{aligned}$$

since $A_{ij}^{(t)}$ is Bernoulli with probability x_iX_j.

Similarly, detailed calculation for the variance is as following:

$$\begin{aligned}
&Cov\left((\bar{A}_{ij} - x_i^TX_j)X_j\right) \\
=&Cov\left(\bar{A}_{ij}X_j - (x_i^TX_j)X_j\right) \\
=&E\left[(\bar{A}_{ij} - x_i^TX_j)^2X_jX_j^T\right] \\
=&E\left[(\bar{A}_{ij}^2 - 2\bar{A}_{ij}x_i^TX_j + (x_i^TX_j)^2)X_jX_j^T\right] \\
=&E\left[E\left[(\bar{A}_{ij}^2 - 2\bar{A}_{ij}x_i^TX_j + (x_i^TX_j)^2)X_jX_j^T|X_j\right]\right] \\
=&E\left[E[\bar{A}_{ij}^2]X_jX_j^T - 2(x_i^TX_j)X_jX_j^TE[\bar{A}_{ij}] + (x_i^TX_j)^2X_jX_j^T\right] \\
=&E\left[(Var(\bar{A}_{ij}) + E[\bar{A}_{ij}]^2)X_jX_j^T - 2(x_i^TX_j)^2X_jX_j^T + (x_i^TX_j)^2X_jX_j^T\right] \\
=&E\left[(\frac{1}{m}x_i^TX_j(1 - x_i^TX_j) + (x_i^TX_j)^2)X_jX_j^T - 2(x_i^TX_j)^2X_jX_j^T + (x_i^TX_j)^2X_jX_j^T\right] \\
=&\frac{1}{m}E\left[X_jX_j^T(x_i^TX_j(1 - x_i^TX_j))\right].
\end{aligned}$$

Therefore, by the classical multivariate central limit theorem and Slutsky's theorem in the multivariate setting, we have

$$\left(\frac{1}{\sqrt{n}} \sum_{j \neq i} (\bar{A}_{ij} - x_i^T X_j) X_j - \frac{x_i^T x_i}{\sqrt{n}} x_i \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \tilde{\Sigma}(x_i)).$$

The strong law of large numbers ensures that $nS^{-1} = n(\tilde{X}^T \tilde{X})^{-1} \to \Delta^{-1} = (E(X_1 X_1^T))^{-1}$ almost surely. In addition,

$$\frac{1}{n} X^T X = \frac{1}{n} W_n \tilde{X}^T \tilde{X} W_n^T = W_n (S/n) W_n^T \to W_n \Delta W_n^T$$

almost surely. Thus we have $W_n \to I$ almost surely. Hence, by Slutsky's theorem in the multivariate setting, we have, conditional on $X_i = x_i$, that

$$\sqrt{n}(Y_i - \tilde{X}_i) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Delta^{-1} \tilde{\Sigma}(x_i) \Delta^{-1}) = \mathcal{N}(0, \bar{\Sigma}(x_i)).$$

The rest are the same. ∎