Web-based Supplementary Materials for "Asymptotic conditional singular value decomposition for high-dimensional genomic data"

Jeffrey T. Leek

Johns Hopkins Bloomberg School of Public Health

Baltimore, MD 21205-2179

jleek@jhsph.edu

Web Appendix A

Proof of Theorem 1

$\boldsymbol{W}^m$ can be broken into five terms:

$$\begin{aligned}\boldsymbol{W}^m &= \frac{1}{m}\boldsymbol{X}^{mT}\boldsymbol{X}^m - \hat{\sigma}^2_{ave}I \\ &= \underbrace{\frac{1}{m}\boldsymbol{G}^T\boldsymbol{\Gamma}^{mT}\boldsymbol{\Gamma}^m\boldsymbol{G}}_{i} + \underbrace{\frac{1}{m}\boldsymbol{G}^T\boldsymbol{\Gamma}^{mT}\boldsymbol{U}^m}_{ii} + \underbrace{\frac{1}{m}\boldsymbol{U}^{mT}\boldsymbol{\Gamma}^m\boldsymbol{G}}_{iii} + \underbrace{\frac{1}{m}\boldsymbol{U}^{mT}\boldsymbol{U}^m}_{iv} - \underbrace{\hat{\sigma}^2_{ave}\boldsymbol{I}}_{v} \quad \text{(A.1)}\end{aligned}$$

We consider each of these terms individually.

i: This term converges to $\boldsymbol{G}^T\boldsymbol{\Delta G}$ by assumption 3.

ii: Let $\boldsymbol{M} = \frac{1}{m}\boldsymbol{G}^T\boldsymbol{\Gamma}^{mT}\boldsymbol{U}^m = \frac{1}{m}\boldsymbol{B}^m\boldsymbol{U}^m$, then $m_{ij} = \frac{1}{m}\sum_{\ell=1}^m b_{i\ell}u_{\ell j}$ where $\mathrm{E}(b_{i\ell}u_{\ell j}) = 0$ and $\mathrm{var}(b_{i\ell}u_{\ell j}) = b_{i\ell}^2\sigma_\ell^2$. So by the Kolmogorov Strong Law of Large Numbers (KSLLN) (Feller, 1968) $m_{ij} \rightarrow_{a.s.} 0$ for all i, j.

iii : By symmetry, this term also converges almost surely to zero.

iv : Let $\boldsymbol{S} = \frac{1}{m}\boldsymbol{U}^{mT}\boldsymbol{U}^m$, and consider the off-diagonal element $s_{ij} = \frac{1}{m}\sum_{\ell=1}^m u_{\ell i}u_{\ell j}$, where $\mathrm{E}(u_{\ell i}u_{\ell j}) = 0$ and $\mathrm{var}(u_{\ell i}u_{\ell j}) = \mathrm{E}(u_{\ell i}^2u_{\ell j}^2) - \mathrm{E}(u_{\ell i}u_{\ell j})^2 = (\sigma_\ell^2)^2$. So again by KSLLN $s_{ij} \rightarrow_{a.s.} 0$. Now consider the diagonal elements $s_{ii} = \frac{1}{m}\sum_{\ell=1}^m u_{\ell i}^2$, where $\mathrm{E}(u_{\ell i}^2) = \sigma_\ell^2$ and $\mathrm{var}(u_{\ell i}^2) = \mathrm{E}(u_{\ell i}^4) - \mathrm{E}(u_{\ell i}^2)^2$. By assumptions 1 the variances are bounded, so by KSLLN $s_{ii} \rightarrow_{a.s.} \bar{\sigma}^2 = \lim_{m\rightarrow\infty}\frac{1}{m}\sum_{i=1}^m \sigma_i^2$, which exists because σ_i^2 is bounded for all i.

Combining terms (i-iv) and applying Slutsky's theorem yields: $\frac{1}{m}\boldsymbol{X}^{mT}\boldsymbol{X}^m \rightarrow_{a.s.} \boldsymbol{G}^T\boldsymbol{\Delta G} +$

$\bar{\sigma}^2\boldsymbol{I}$. Since the eigenvalues of a matrix are defined as roots of a determinant depending on the elements of that matrix, and since the roots of a polynomial equation are a continuous multi-valued function of the coefficients (Henriksen and Isbell, 1953), the eigenvalue function is continuous. The eigenvalues of $\frac{1}{m}\boldsymbol{X}^{mT}\boldsymbol{X}^m$ converge almost surely to the eigenvalues of $\boldsymbol{G}^T\boldsymbol{\Delta G}+\bar{\sigma}^2\boldsymbol{I}$ from the continuous mapping theorem. The eigenvalues of $\boldsymbol{G}^T\boldsymbol{\Delta G}+\bar{\sigma}^2\boldsymbol{I}$ are equal to $\lambda_1+\bar{\sigma}^2,\ldots,\lambda_n+\bar{\sigma}^2$; but $\lambda_{r+1},\ldots,\lambda_n$ are equal to zero by assumption, so the last $n-r$ eigenvalues consistently estimate $\bar{\sigma}^2$.

v : $\hat{\sigma}^2_{ave} = \frac{1}{m}\sum_{i=1}^m \frac{1}{(n-\kappa)}\sum_{j=1}^n \left(x_{ij}-\sum_{k=1}^{\kappa}\hat{\gamma}_{ik}\hat{v}_{kj}\right)^2 = \frac{1}{n-\kappa}\sum_{k=\kappa}^n \lambda_k(\boldsymbol{Z}^m)$, where $\lambda_k(\boldsymbol{Z}^m)$ is the kth eigenvalue of $\boldsymbol{Z}_m = \frac{1}{m}\boldsymbol{X}^{mT}\boldsymbol{X}^m$. But for all $\kappa > r$, $\lambda_k(\boldsymbol{Z}^m)k$ converges to $\bar{\sigma}^2$ almost surely.

Combining terms (i-v) and applying Slutsky's theorem yields: $\boldsymbol{W}^m \rightarrow_{a.s.} \boldsymbol{G}^T\boldsymbol{\Delta G}$. By the same argument as above, the eigenvalues of $\boldsymbol{W}^m$ converge almost surely to the eigenvalues of $\boldsymbol{G}^T\boldsymbol{\Delta G}$ from the continuous mapping theorem. Further, since both the matrix W^m and the eigenvalues converge almost surely, and the eigenvectors can be obtained from a linear operation of these two elements, the eigenvectors of $\boldsymbol{W}^m$ corresponding to the unique eigenvalues must converge to the corresponding eigenvectors of $\boldsymbol{G}^T\boldsymbol{\Delta G}$.

Proof of Lemma 1

We wish to show that the indicator:

$$1\left\{\lambda_k(\boldsymbol{W}^m) \geq c_m\right\} = 1\left\{\frac{1}{c_m}\lambda_k(\boldsymbol{W}^m) \geq 1\right\}.$$

consistently distinguishes between zero and non-zero eigenvalues. If $\lambda_k(\boldsymbol{W}^m) = \lambda_k + O_P\left(m^{-\frac{1}{2}}\right)$ then for any $c_m = O(m^{-\eta})$, $0<\eta<\frac{1}{2}$, when $\lambda_k = 0$, $\left(\frac{1}{c_m}\lambda_k(\boldsymbol{W}^m)\right) = O(m^{\eta})O_P\left(m^{-\frac{1}{2}}\right) = m^{\eta-\frac{1}{2}}O_P(1)$ and $1\left\{\frac{1}{c_m}\lambda_k(\boldsymbol{W}^m)\geq 1\right\} \rightarrow_P 0$. When $\lambda_k > 0$, $\left(\frac{1}{c_m}\lambda_k(\boldsymbol{W}^m)\right) = O(m^{\eta})\left\{\lambda_k + O_P(m^{-\frac{1}{2}})\right\} = O(m^{\eta}) + o_P\left\{m^{\eta-\frac{1}{2}}\right\} \rightarrow \infty$. So when $\lambda_k > 0$ and $0<\eta<\frac{1}{2}$, $1\left\{\frac{1}{c_m}\lambda_k(\boldsymbol{W}^m)\geq 1\right\} \rightarrow_P 1$. To complete the proof, we must show that $\lambda_k(\boldsymbol{W}^m) = \lambda_k + O_P\left(m^{-\frac{1}{2}}\right)$.

From the decomposition (A.1) we can write $\boldsymbol{W}^m$ as a continuous function of:

$$\boldsymbol{y}_i = \frac{1}{m}\left(k_{i1}\boldsymbol{u}_{i.}^T, \ldots, k_{in}\boldsymbol{u}_{i.}^T, u_{i1}\boldsymbol{u}_{i.}^T, \ldots, u_{in}\boldsymbol{u}_{i.}^T\right)^T,$$

and $\frac{1}{m}\boldsymbol{G}^T\boldsymbol{\Gamma}^{mT}\boldsymbol{\Gamma}^m\boldsymbol{G}$, where $\boldsymbol{K} = \boldsymbol{G}^T\boldsymbol{\Gamma}^{mT}$, this is straightforward for components $(i - iv)$, for component v:

$$\hat{\sigma}^2_{ave} = \frac{1}{n-\kappa}\sum_{k=\kappa}^{n}\lambda_k(\boldsymbol{Z}^m)$$

but $\boldsymbol{Z}^m = \frac{1}{m}\boldsymbol{X}^{mT}\boldsymbol{X}^m$ is a continuous function of $\boldsymbol{y}_i$ and $\frac{1}{m}\boldsymbol{G}^T\boldsymbol{\Gamma}^{mT}\boldsymbol{\Gamma}^m\boldsymbol{G}$ and the eigenvalues of $\boldsymbol{Z}^m$ are a continuous function of the matrix (Henriksen and Isbell, 1953), so $\hat{\sigma}^2_{ave}$ is a continuous function of $\boldsymbol{y}_i$.

The expectation of $\boldsymbol{y}_i$ is $\mathrm{E}(\boldsymbol{y}_i) = (0, \ldots, 0, \sigma_i^2, 0, \ldots, 0, \sigma_i^2, 0, \ldots, 0, \sigma_i^2)^T$. Define $\boldsymbol{y}_i^* = \sqrt{m}\{\boldsymbol{y}_i - \mathrm{E}(\boldsymbol{y}_i)\}$; the covariance matrix for this random variable is $\mathrm{cov}(\boldsymbol{y}_i^*) = \frac{1}{m}\boldsymbol{\Sigma}_i$. From assumption 1, $\frac{1}{m}\sum_{i=1}^m \boldsymbol{\Sigma}_i \to \boldsymbol{\Sigma}$ and $\sum_{i=1}^m \boldsymbol{y}_i^*$ is asymptotically normally distributed if the Lindeberg condition holds for every $\epsilon > 0$. Let

$$\psi_i = \left\{\sum_{j=1}^n (u_{ij}^2 - \sigma_i^2)^2 + \sum_{j=1}^n\sum_{k=1}^n k_{ij}^2 u_{ik}^2 + \sum_{k\neq j} u_{ij}^2 u_{ik}^2\right\}\boldsymbol{1}(\parallel \boldsymbol{y}_i^* \parallel^2 > \epsilon).$$

The Lindeberg condition requires $\mathrm{E}(\psi_i) \to 0$ for every i. But ψ_i is only non-zero when:

$$\parallel \boldsymbol{y}_i^* \parallel^2 = \frac{1}{m}\left\{\sum_{j=1}^n (u_{ij}^2 - \sigma_i^2)^2 + \sum_{j=1}^n\sum_{k=1}^n k_{ij}^2 u_{ik}^2 + \sum_{k\neq j} u_{ij}^2 u_{ik}^2\right\} > \epsilon$$

an event that has probability zero as $m \to \infty$, so $\psi_i \to_P 0$. It is also clear that $|\psi_i| \leq m \parallel \boldsymbol{y}_i^* \parallel^2$ and $\mathrm{E}\{m \parallel \boldsymbol{y}_i^* \parallel^2\} < \infty$ by assumption 1. So by the dominated convergence theorem $\mathrm{E}(\psi_i) \to_P 0$ for each i and hence for every $\epsilon > 0$,

$$\sum_{i=1}^m \mathrm{E}\{\parallel \boldsymbol{y}_i^* \parallel \boldsymbol{1}(\parallel \boldsymbol{y}_i^* \parallel > \epsilon)\} = \frac{1}{m}\sum_{i=1}^m \mathrm{E}\{\psi_i\} \to_P 0$$

Since the Lindeberg condition is satisfied $\sum_{i=1}^m \boldsymbol{y}_i^*$ is asymptotically normally distributed. Since $\mathrm{vec}(\boldsymbol{W}^m) = \boldsymbol{g}(\sum_{i=1}^m \boldsymbol{y}_i) + \mathrm{vec}\left(\frac{1}{m}\boldsymbol{G}\boldsymbol{\Gamma}^{mT}\boldsymbol{\Gamma}^m\boldsymbol{G}\right)$, where the function vec(.) concatenates the columns of a matrix and $\boldsymbol{g}$ is a continuous function,

$$\sqrt{m}\left(\mathrm{vec}(\boldsymbol{W}^m) - \mathrm{vec}\left(\frac{1}{m}\boldsymbol{G}^T\boldsymbol{\Gamma}^{mT}\boldsymbol{\Gamma}^m\boldsymbol{G}\right)\right) \to \mathrm{MVN}(\boldsymbol{0}, \boldsymbol{\Sigma}_w)$$

by the multivariate delta method, so $\sqrt{m}(\boldsymbol{W}^m - \mathrm{vec}\left(\frac{1}{m}\boldsymbol{G}^T\boldsymbol{\Gamma}^{mT}\boldsymbol{\Gamma}^m\boldsymbol{G}\right)) = O_P(1)$. Since $\lambda_r - \lambda_{r+1} = c > 0$ and $\boldsymbol{W}^m$ is symmetric and real, by Theorem 4.2 of Eaton and Tyler (1991),

$$\begin{aligned} \sqrt{m}\left\{\lambda_1(\boldsymbol{W}^m), \ldots, \lambda_n(\boldsymbol{W}^m))^T - (\lambda_1, \ldots, \lambda_n)^T\right\} &= O_P(1) \\ \Rightarrow \sqrt{m}\lambda_k(\boldsymbol{W}^m) &= \sqrt{m}\lambda_k + O_P(1) \quad \forall k. \end{aligned}$$

So $\lambda_k(\boldsymbol{W}^m) = \lambda_k + O_P(m^{-1/2})$, which completes the proof.

Proof of Corollary 1

$$\begin{aligned} \boldsymbol{R}^m &= \boldsymbol{X}^m(\boldsymbol{I} - \boldsymbol{S}(\boldsymbol{S}^T\boldsymbol{S})^{-1}\boldsymbol{S}^T) \\ &= \{\boldsymbol{B}^m\boldsymbol{S} + \boldsymbol{\Gamma}^m\boldsymbol{G} + \boldsymbol{U}^m\}\{\boldsymbol{I} - \boldsymbol{S}^T(\boldsymbol{S}\boldsymbol{S}^T)^{-1}\boldsymbol{S}\} \\ &= \boldsymbol{\Gamma}^m\boldsymbol{G} + \boldsymbol{U}^m\boldsymbol{P}_s \end{aligned}$$

Then we can write:

$$\begin{aligned} \boldsymbol{W}_R^m &= \frac{1}{m}\boldsymbol{R}^{mT}\boldsymbol{R}^m - \hat{\sigma}^2_{ave}I \\ &= \underbrace{\frac{1}{m}\boldsymbol{G}^T\boldsymbol{\Gamma}^{mT}\boldsymbol{\Gamma}^m\boldsymbol{G}}_{i} + \underbrace{\frac{1}{m}\boldsymbol{G}^T\boldsymbol{\Gamma}^{mT}\boldsymbol{U}\boldsymbol{P}_S}_{ii} + \underbrace{\frac{1}{m}\boldsymbol{P}_s^T\boldsymbol{U}^{mT}\boldsymbol{\Gamma}^m\boldsymbol{G}}_{iii} + \underbrace{\frac{1}{m}\boldsymbol{P}_s^T\boldsymbol{U}^{mT}\boldsymbol{U}^m\boldsymbol{P}_s}_{iv} - \underbrace{\boldsymbol{P}_s^T\hat{\sigma}^2_{ave}\boldsymbol{P}_s}_{v} \end{aligned}$$

Following the proof of Theorem 1, terms (ii) and (iii) converge to zero, term (i) converges to $\boldsymbol{G}^T\boldsymbol{\Delta}\boldsymbol{G}$, and term (iv) converges to $\boldsymbol{P}_s^T\bar{\sigma}^2\boldsymbol{P}_s = \bar{\sigma}^2\boldsymbol{P}_s$. Term (v) is equal to $\widehat{\sigma}^2_{ave}\boldsymbol{P}_s$, but $\widehat{\sigma}^2_{ave}$ is equal to $\frac{1}{n-d-k}\sum_{k=(d+\kappa)}^{n}\lambda_k(\boldsymbol{Z}^m)$, which converges almost surely to $\bar{\sigma}^2$. Thus, $\boldsymbol{W}_{\boldsymbol{R}}^m$

converges almost surely to $\boldsymbol{G}^T \boldsymbol{\Delta} \boldsymbol{G}$ and the result follows according to the proof of Theorem 1.

Proof of Corollary 2

The proof follows the proof of Lemma 1, where the function, $\boldsymbol{g}$, incorporates the project term $\boldsymbol{P}_s$.

Web Appendix B

[Figure 1 about here.]

[Figure 2 about here.]

[Figure 3 about here.]

[Figure 4 about here.]

[Figure 5 about here.]

[Figure 6 about here.]

[Figure 7 about here.]

[Figure 8 about here.]

Web Appendix C

[Table 1 about here.]

References

Eaton, M. L. and Tyler, D. E. (1991). On wielandt's inequality and its application to the asymptotic distribution of the eigenvalues of a random symmetric matrix. *Ann Stat* **19,** 260–271.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley, 3 edition.

Henriksen, M. and Isbell, J. R. (1953). On the continuity of the real roots of an algebraic equation. *Proc Amer Math Soc* **4,** 431–4.

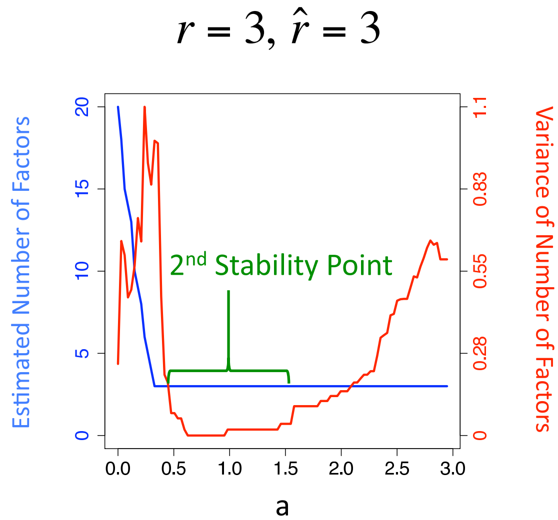

Figure 1. A plot of the estimated number of factors (blue and left axis) and the empirical variance of the estimate for varying set sizes (red and right axis) across a range of coefficients a for a simulated example with $r = 3$. The second stability point (green bracket) is the second point, moving from left to right, where the variance finds a trough. Hallin & Liska (2007) suggest using the estimate corresponding to this second stability point as a practical estimator of the number of factors.

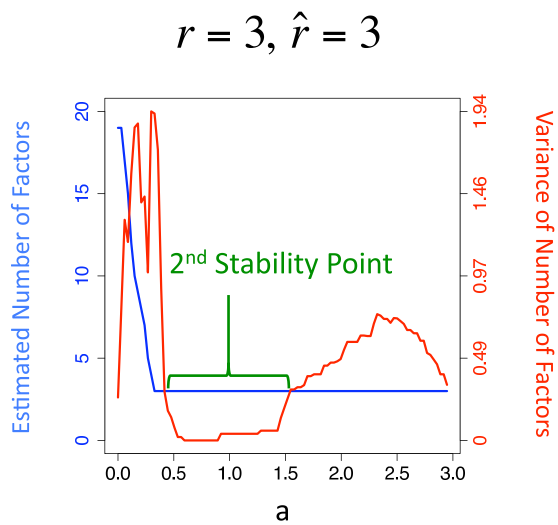

Figure 2. A plot of the estimated number of factors (blue and left axis) and the empirical variance of the estimate for varying set sizes (red and right axis) across a range of coefficients a for a simulated example with $r = 3$. The second stability point (green bracket) is the second point, moving from left to right, where the variance finds a trough. Hallin & Liska (2007) suggest using the estimate corresponding to this second stability point as a practical estimator of the number of factors.

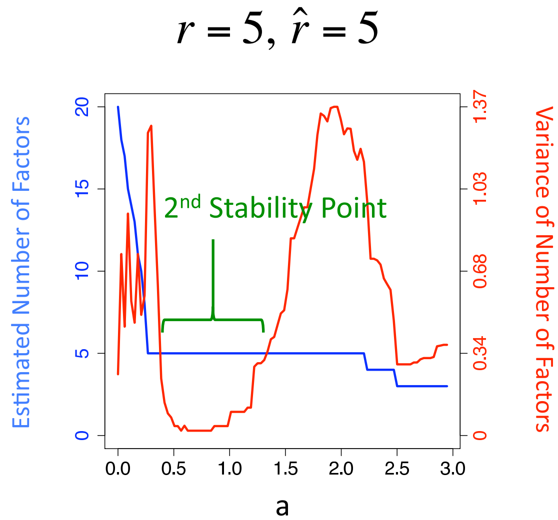

Figure 3. A plot of the estimated number of factors (blue and left axis) and the empirical variance of the estimate for varying set sizes (red and right axis) across a range of coefficients a for a simulated example with $r = 5$. The second stability point (green bracket) is the second point, moving from left to right, where the variance finds a trough. Hallin & Liska (2007) suggest using the estimate corresponding to this second stability point as a practical estimator of the number of factors.

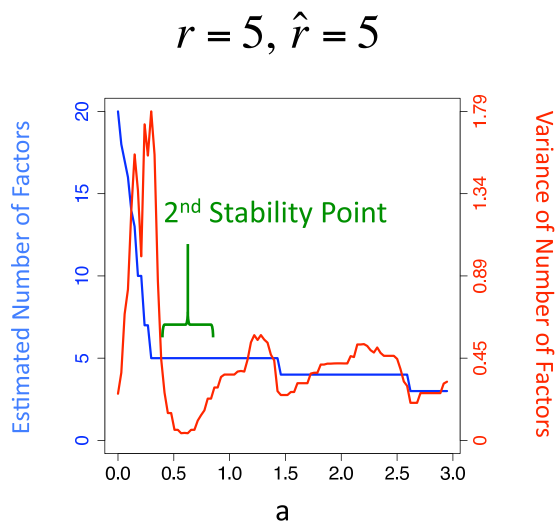

Figure 4. A plot of the estimated number of factors (blue and left axis) and the empirical variance of the estimate for varying set sizes (red and right axis) across a range of coefficients a for a simulated example with $r = 5$. The second stability point (green bracket) is the second point, moving from left to right, where the variance finds a trough. Hallin & Liska (2007) suggest using the estimate corresponding to this second stability point as a practical estimator of the number of factors.

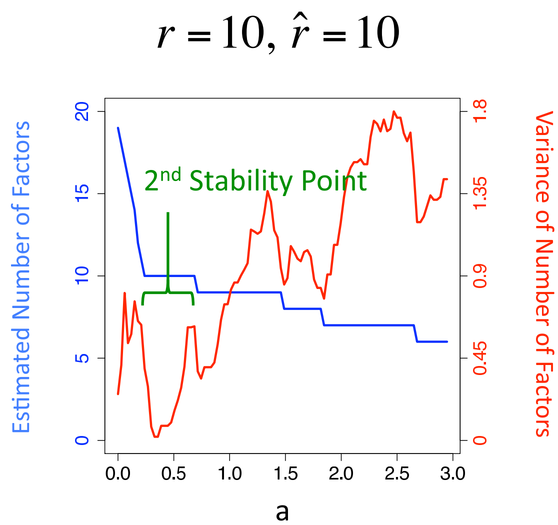

Figure 5. A plot of the estimated number of factors (blue and left axis) and the empirical variance of the estimate for varying set sizes (red and right axis) across a range of coefficients a for a simulated example with $r = 10$. The second stability point (green bracket) is the second point, moving from left to right, where the variance finds a trough. Hallin & Liska (2007) suggest using the estimate corresponding to this second stability point as a practical estimator of the number of factors.

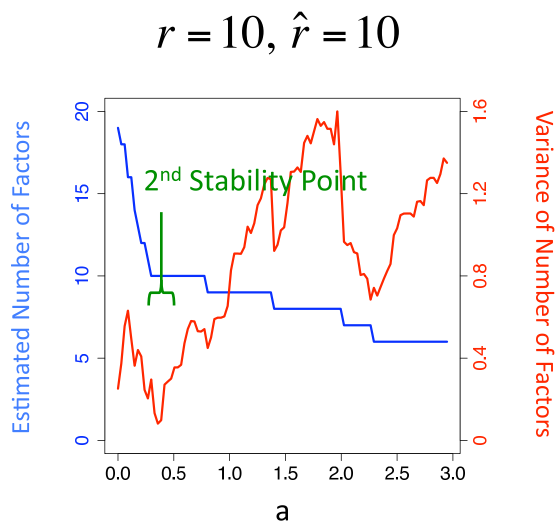

Figure 6. A plot of the estimated number of factors (blue and left axis) and the empirical variance of the estimate for varying set sizes (red and right axis) across a range of coefficients a for a simulated example with $r = 10$. The second stability point (green bracket) is the second point, moving from left to right, where the variance finds a trough. Hallin & Liska (2007) suggest using the estimate corresponding to this second stability point as a practical estimator of the number of factors.

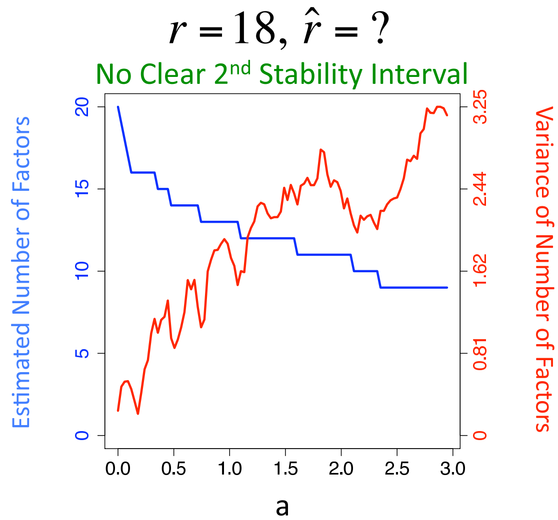

Figure 7. A plot of the estimated number of factors (blue and left axis) and the empirical variance of the estimate for varying set sizes (red and right axis) across a range of coefficients a for a simulated example with $r = 18$. The second stability point (green bracket) is the second point, moving from left to right, where the variance finds a trough. Hallin & Liska (2007) suggest using the estimate corresponding to this second stability point as a practical estimator of the number of factors. Since $r = 18$ is close to the sample size $n = 20$, there is no clear second stability point, so the Hallin & Liska approach does not give an estimate.

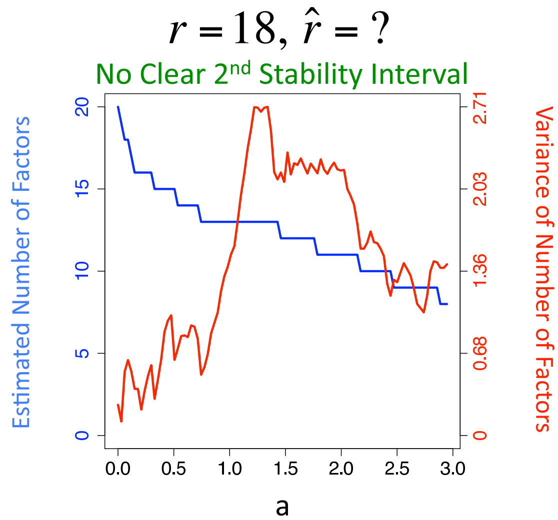

Figure 8. A plot of the estimated number of factors (blue and left axis) and the empirical variance of the estimate for varying set sizes (red and right axis) across a range of coefficients a for a simulated example with $r = 18$. The second stability point (green bracket) is the second point, moving from left to right, where the variance finds a trough. Hallin & Liska (2007) suggest using the estimate corresponding to this second stability point as a practical estimator of the number of factors. Since $r = 18$ is close to the sample size $n = 20$, there is no clear second stability point, so the Hallin & Liska approach does not give an estimate.

Table 1

Results from a simulation experiment. For each combination of m, n and r, 100 independent microarray data sets were simulated. The average (s.d.) RMSFE, a measure of how well the eigenvectors of W_m span the linear space spanned by G, is reported for the Lemma 1 estimator of r and the Bai & Ng (2002) and Buja & Eyuboglu (1992) estimators.

(m, n)	r	RMSFE($\widehat{r}$) $\times 10^5$	RMSFE($\widehat{r}_{bn}$) $\times 10^5$	RMSFE($\widehat{r}_{be}$) $\times 10^5$
(1000,10)	3	403.84 (773.88)	2514.19 (2881.21)	915.18 (1571.98)
(5000,10)	3	78.44 (271.25)	3022.17 (2409.22)	785.95 (1446.20)
(10000,10)	3	50.18 (215.60)	2881.82 (2585.60)	817.71 (1753.53)
(1000,20)	3	109.77 (23.18)	685.79 (1683.42)	133.18 (237.64)
(5000,20)	3	22.49 (4.67)	426.71 (1266.14)	22.49 (4.67)
(10000,20)	3	10.42 (2.22)	193.79 (972.41)	10.42 (2.22)
(1000,100)	3	101.24 (7.29)	101.24 (7.29)	101.24 (7.29)
(5000,100)	3	20.26 (1.57)	20.26 (1.57)	20.26 (1.57)
(10000,100)	3	10.11(0.82)	10.11(0.82)	10.11 (0.82)
(1000,10)	5	822.46 (624.17)	2514 .19 (1704.76)	3256.58 (1698.78)
(5000,10)	5	224.48(309.48)	3022.17 (1729.20)	2885.43 (1636.53)
(10000,10)	5	129.70(245.14)	2606.50 (1371.87)	2655.60 (1474.46)
(1000,20)	5	395.34 (591.32)	1378.21 (1101.71)	480.67 (718.98)
(5000,20)	5	46.95 (151.17)	1298.78 (1109.70)	230.22 (563.62)
(10000,20)	5	18.00 (78.65)	980.16 (991.06)	191.60 (478.68)
(1000,100)	5	99.65 (7.22)	99.65 (7.22)	99.65 (7.21)
(5000,100)	5	20.11(1.31)	20.11 (1.31)	20.11 (1.31)
(10000,100)	5	10.10 (0.63)	10.10 (0.63)	10.10 (0.63)
(1000,20)	10	1108.28 (435.75)	1732.07 (644.47)	3034.82 (1076.86)
(5000,20)	10	297.95 (245.09)	1715.14 (545.61)	2683.19 (717.63)
(10000,20)	10	198.12 (164.99)	1592.15 (520.36)	2562.26 (711.42)
(1000,100)	10	96.40 (5.31)	96.40 (5.31)	96.40 (5.31)
(5000,100)	10	19.30 (1.06)	19.30 (1.06)	19.30 (1.06)
(10000,100)	10	9.60 (0.44)	9.60(0.44)	9.60 (0.44)
(1000,20)	18	1148.48 (273.08)	1562.41(329.37)	5479.46 (1011.77)
(5000,20)	18	497.17(144.31)	1484.74 (312.31)	5033.88 (832.27)
(10000,20)	18	336.51 (104.66)	1472.69 (350.24)	4876.27 (867.75)
(1000,100)	18	314.30 (293.42)	90.25 (3.61)	97.64 (53.06)
(5000,100)	18	17.88 (0.74)	17.88 (0.74)	17.88 (0.74)
(10000,100)	18	8.88 (0.34)	8.88 (0.34)	8.88 (0.34)