Supplemental materials for

Adaptive Minimal Confidence Region Rule for Multivariate Initialization Bias Truncation in Discrete-event Simulations

Jianguo Wu, Honglun Xu, Feng Ju and Bill Tseng

August 30, 2019

1. Proof of Theorem 1

Thm1.1. From Eq. (12) we can get

$$E(R_n) = \frac{\prod_{i=1}^{p}(n-i)}{n^h(n-1)^p}|\Sigma|$$

It is straightforward that $\lim_{n\to\infty} E(R_n) = 0$ for $h > 0$. The ratio $E(R_{n+1})/E(R_n)$ can be derived as

$$f_1(n) = \frac{E(R_{n+1})}{E(R_n)} = \frac{\dfrac{\prod_{k=1}^{p}(n+1-k)}{(n+1)^h n^p}}{\dfrac{\prod_{k=1}^{p}(n-k)}{n^h(n-1)^p}} = \frac{n^{h-p+1}}{(n-p)}\frac{(n-1)^p}{(n+1)^h}$$

The derivative of the ratio with respect to n can be obtained as

$$f_1'(n) = \frac{n^{h-p}(n-1)^{p-1}(n+1)^{h-1}}{[(n-p)(n+1)^h]^2}[hn^2 - n(p^2+ph+h-p) + p(h-p+1)]$$

Therefore, when $n > g_1(p,h) = \frac{(p^2+ph+h-p)+\sqrt{(p^2+ph+h-p)^2-4ph(h-p+1)}}{2h}$, $f_1'(n) > 0$, $f_1(n)$ increases monotonically.

Since $\lim_{n\to\infty} f_1(n) = 1$, we can get $\frac{E(R_{n+1})}{E(R_n)} < 1$ or $E(R_{n+1}) < E(R_n)$ for $g_1(p,h) < n < \infty$. Therefore $E(R_n)$ decreases monotonically for $n > g_1(p,h)$.

Thm1.2. From Eq. (12) we can get

$$Var\left(\frac{|S_{1:n}|}{n^h}\right) = \frac{\prod_{i=1}^{p}(n-i)\left[\prod_{j=1}^{p}(n-j+2) - \prod_{j=1}^{p}(n-j)\right]}{n^{2h}(n-1)^{2p}}|\Sigma|^2$$

Similarly we can obtain the ratio

$$f_2(n) = \frac{Var\left(\frac{|S_{1:n+1}|}{(n+1)^h}\right)}{Var\left(\frac{|S_{1:n}|}{n^h}\right)} = \frac{(n-1)^{2p}n^{2h-2p+2}(2n-h+3)}{(n+1)^{2h}(n-p)(n-p+2)(2n-p+1)}$$

Following the same approach, we can show that when $n > g_2(p,h)$ where g_2 is the largest root of equation $f_2'(n) = 0$, $f_2(n)$ increases monotonically. Since $\lim_{n\to\infty} f_2(n) = 1$, we can get $Var\left(\frac{|S_{1:n+1}|}{(n+1)^h}\right) < Var\left(\frac{|S_{1:n}|}{n^h}\right)$ for $g_2(p,h) < n < \infty$. Therefore $Var\left(\frac{|S_{1:n}|}{n^h}\right)$ decreases monotonically for $n > g_2(p,h)$. Note that the closed form of $g_2(p,h)$ is very difficult to obtain, if not impossible. We can numerically calculate it instead. Alternatively, the variance can be reformulated as

$$Var\left(\frac{|S_{1:n}|}{n^h}\right) = \frac{2p|\Sigma|^2}{(n-1)n^{2h}}\left[\left(1-\frac{p-2}{n-1}\right)\left(1-\frac{p-1}{n-1}\right)(1-\frac{p-3}{2(n-1)})\right]\prod_{i=1}^{p-3}\left(1-\frac{i}{n-1}\right)^2$$

Since for sufficiently large n,

$$\left[\left(1-\frac{p-2}{n-1}\right)\left(1-\frac{p-1}{n-1}\right)(1-\frac{p-3}{2(n-1)})\right]\prod_{i=1}^{p-3}\left(1-\frac{i}{n-1}\right)^2 \to 1,$$

and $\frac{2p}{(n-1)n^h}$ monotonically decreases, there must exist $g_2(p,h)$ such that when $n > g_2(p,h)$, $Var\left(\frac{|S_{1:n}|}{n^h}\right)$ is monotonically decreasing. Clearly $\lim_{n\to\infty} Var\left(\frac{|S_{1:n}|}{n^h}\right) = 0$.

From the derivation process we can clearly see that $f_1(n)$ and $f_2(n)$ may also be less than 1 for certain n with $n < g_1$ and $n < g_2$ respectively, which means g_1 and g_2 are just upper bounds for the lower ends of the decreasing regions.

2. Proof of Theorem 2

Thm 2.1. The non-centrality matrix $T = \sum_{i=1}^{n+1}(\mu_i - \bar{\mu})(\mu_i - \bar{\mu})^T$ where $\mu_i = \mu_0$ for $i = 1, \dots, n$, $\mu_{n+1} = \mu$, and $\bar{\mu} = \frac{\sum_{i=1}^{n+1}\mu_i}{n+1}$.

$$|T - k^2\Sigma| = 0 => \left|\Sigma^{-\frac{1}{2}}T\Sigma^{-\frac{1}{2}} - k^2 I\right| = 0$$

$$=> (n+1)^p \left|\frac{1}{\mathrm{n}+1}\mathrm{ZZ}^{\mathrm{T}} - \frac{k^2}{n+1}I\right| = 0$$

where

$$Z = \left(\Sigma^{-\frac{1}{2}}(\mu_0 - \bar{\mu}), \Sigma^{-\frac{1}{2}}(\mu_0 - \bar{\mu}), \dots, \Sigma^{-\frac{1}{2}}(\mu - \bar{\mu})\right)$$

Therefore $\frac{k_i^2}{n+1}, i = 1, \dots, p$ are the eigenvalues of the sample covariance matrix $\frac{1}{\mathrm{n}+1}ZZ^{\mathrm{T}}$. Based on the theory of principal component analysis (PCA) (Izenman 2008), it is straightforward that there is only one non-zero eigenvalue since all samples $\left(\Sigma^{-\frac{1}{2}}(\mu_0 - \bar{\mu}), \Sigma^{-\frac{1}{2}}(\mu_0 - \bar{\mu}), \dots, \Sigma^{-\frac{1}{2}}(\mu - \bar{\mu})\right)$ are on a line. The corresponding eigenvector is

$$\alpha_1 = \frac{\Sigma^{-\frac{1}{2}}(\mu - \mu_0)}{\left|\Sigma^{-\frac{1}{2}}(\mu - \mu_0)\right|}$$

Therefore the largest eigenvalue, which is the sample variance of the first principal components, can be computed as

$$\frac{k_1^2}{n+1} = \alpha_1^T \frac{1}{\mathrm{n}+1}\mathrm{ZZ}^{\mathrm{T}}\alpha_1$$

$$k_1^2 = \alpha_1^T \mathrm{ZZ}^{\mathrm{T}}\alpha_1 = \sum_{i=1}^{n+1}\left[\alpha_1^T\Sigma^{-\frac{1}{2}}(\mu_i - \bar{\mu})\right]^2$$

$$= \frac{n\left[(\mu - \mu_0)^T\Sigma^{-1}(\mu - \mu_0)\right]^2}{(n+1)^2(\mu - \mu_0)^T\Sigma^{-1}(\mu - \mu_0)} + \frac{n^2\left[(\mu - \mu_0)^T\Sigma^{-1}(\mu - \mu_0)\right]^2}{(n+1)^2(\mu - \mu_0)^T\Sigma^{-1}(\mu - \mu_0)}$$

$$= \frac{n}{n+1}(\mu - \mu_0)^T\Sigma^{-1}(\mu - \mu_0)$$

Thm 2.2. $S_{1:n+1}$ follows noncentral Wishart distribution of degree n (Anderson and Mathématicien 1958). Since all $\mu_i, i = 1, \ldots, n+1$ are on a line (noncentral linear case), the expectation of the determinant of the generalized variance is tractable. According to Anderson (Anderson 1946), the expectation of the mth moment of $|S_{1:n+1}|$ can be calculated as follows:

$$E(|S_{1:n+1}|^m) = \frac{2^{pm}}{n^{pm}} |\Sigma|^m \exp(-\frac{1}{2}k_1^2) \prod_{i=1}^{p-1} \frac{\Gamma\left[\frac{1}{2}(n-i)+m\right]}{\Gamma\left[\frac{1}{2}(n-i)\right]} \sum_{j=0}^{\infty} \left\{ \frac{k_1^{2j}}{2^j j!} \frac{\Gamma\left(\frac{1}{2}n+j+m\right)}{\Gamma\left(\frac{1}{2}n+j\right)} \right\}$$

For $m = 1$

$$E(|S_{1:n+1}|) = \frac{2^p}{n^p} |\Sigma| \exp(-\frac{1}{2}k_1^2) \left[\prod_{i=1}^{p-1} \frac{1}{2}(n-i)\right] \sum_{j=0}^{\infty} \frac{k_1^{2j}\left(\frac{1}{2}n+j\right)}{2^j j!}$$

Since

$$\sum_{j=0}^{\infty} \frac{k_1^{2j}\left(\frac{1}{2}n+j\right)}{2^j j!} = \frac{1}{2}n \sum_{j=0}^{\infty} \frac{\left(\frac{k_1^2}{2}\right)^j}{j!} + \sum_{j=1}^{\infty} \frac{\left(\frac{k_1^2}{2}\right)^j}{(j-1)!} = \left(\frac{n}{2} + \frac{k_1^2}{2}\right) \exp\left(\frac{k_1^2}{2}\right),$$

Therefore

$$E(|S_{1:n+1}|) = \frac{|\Sigma|(n+k_1^2)}{n^p(n-p)} \prod_{i=1}^{p} (n-i)$$

3. Proof of Theorem 3

From Eq. (12) and Eq. (14) we obtain that

$$E(R_n) = E\left(\frac{|S_{1:n}|}{n^h}\right) = \frac{\prod_{i=1}^{p}(n-i)}{n^h(n-1)^p} |\Sigma|$$

and

$$E(R_{n+1}) = E\left(\frac{|S_{1:n+1}|}{(n+1)^h}\right) = \frac{|\Sigma|(n+k_1^2)}{(n+1)^h n^p (n-p)} \prod_{i=1}^{p} (n-i)$$

Define $f(n) = E(R_{n+1})/E(R_n)$, then

$$f(n) = \frac{(n-1)^p n^{h-p} (n+k_1^2)}{(n-p)(n+1)^h}$$

Let $f(n) > 1$,

$$\frac{(n-1)^p n^{h-p}(n+k_1^2)}{(n-p)(n+1)^h} > 1$$

$$k_1^2 > \frac{(n-p)(n+1)^h}{(n-1)^p n^{h-p}} - n$$

$$(\mu-\mu')^T\Sigma^{-1}(\mu-\mu') > \frac{(n-p)(n+1)^{h+1}}{(n-1)^p n^{h-p+1}} - (n+1)$$

$$= (n+1)\left[\frac{(n-p)(n+1)^h}{(n-1)^p n^{h-p+1}} - 1\right]$$

For a large n, $\left(1+\frac{1}{n}\right)^h \approx 1+\frac{h}{n}$, $\left(1-\frac{1}{n}\right)^p \approx 1-\frac{p}{n}$, Therefore

$$\frac{(n-p)(n+1)^h}{(n-1)^p n^{h-p+1}} \approx 1+\frac{h}{n}$$

$$(n+1)\left[\frac{(n-p)(n+1)^h}{(n-1)^p n^{h-p+1}} - 1\right] \approx \frac{(n+1)h}{n} \approx h$$

4. Proof of Theorem 4

Thm 4.1. Since all the means $\mu_i, i = c-t+1, \dots, c+n$ are in a linear relationship, or they are on a line, we can use the same approach as was used in Thm 2.1.

$$\bar{\mu} = \frac{\sum_{i=c-t+1}^{c}\left[\mu_b + \frac{(i-1)(\mu_0-\mu_b)}{c}\right] + n\mu_0}{n+t} = \frac{\left[\frac{(2c-t-1)t(\mu_0-\mu_b)}{2c} + t\mu_b + n\mu_0\right]}{n+t}$$

For $i = c-t+1, \dots, c$,

$$\mu_i - \bar{\mu} = \mu_b + \frac{(i-1)(\mu_0-\mu_b)}{c} - \frac{\left[\frac{(2c-t-1)t(\mu_0-\mu_b)}{2c} + t\mu_b + n\mu_0\right]}{n+t}$$

$$= \frac{(\mu_0-\mu_b)}{2c(n+t)}[-2nc + 2(i-1)(n+t) - (2c-t-1)t]$$

For $i = c+1, \dots, c+n$,

$$\mu_i - \bar{\mu} = \mu_0 - \frac{\left[\frac{(2c-t-1)t\Delta\mu}{2c} + t\mu_b + n\mu_0\right]}{n+t} = \frac{(\mu_0-\mu_b)t(t+1)}{2c(n+t)}$$

Therefore

$$k_{1,c-t+1:n+c}^2 = \sum_{i=c-t+1}^{n+c} \left[\alpha_1^T \Sigma^{-\frac{1}{2}} (\mu_i - \bar{\mu})\right]^2$$

$$= \frac{D_s^2}{4c^2(n+t)^2} \left\{ nt^2(t+1)^2 + \sum\nolimits_{i=c-t+1}^{c} [-2nc + 2(i-1)(n+t) - (2c-t-1)t]^2 \right\}$$

Thm 4.2. $S_{c-t+1:n+c}$ follows noncentral Wishart distribution of degree n (Anderson and Mathématicien 1958). Based on Thm 2.2 and Eq. (14) we can easily get

$$E(|S_{c-t+1:n+c}|) = \frac{|\Sigma|(n+t-1+k_{1,c-t+1:n+c}^2)}{(n+t-1)^p(n+t-1-p)} \prod\nolimits_{i=1}^{p} (n+t-1-i)$$

Thm 4.3.

$$E(R_{c-t+1:n+c}) = \frac{|\Sigma|(n+t-1+k_{1,c-t+1:n+c}^2)}{(n+t)^h(n+t-1)^p(n+t-1-p)} \prod\nolimits_{i=1}^{p} (n+t-1-i)$$

Let

$$f(t) = \frac{E(R_{c-t:n+c})}{E(R_{c-t+1:n+c})}$$

We can get

$$f(t) = \frac{\dfrac{\left(n+t+k_{1,c-t:n+c}^2\right)\prod_{i=1}^{p-1}(n+t-i)}{(n+t+1)^h(n+t)^p}}{\dfrac{\left(n+t-1+k_{1,c-t+1:n+c}^2\right)\prod_{i=1}^{p-1}(n+t-1-i)}{(n+t)^h(n+t-1)^p}}$$

$$= \frac{\left(n+t+k_{1,c-t:n+c}^2\right)}{\left(n+t-1+k_{1,c-t+1:n+c}^2\right)} \frac{(n+t-1)^p}{(n+t)^p} \frac{(n+t)^h}{(n+t+1)^h} \frac{n+t-1}{n+t-p}$$

$$\approx \frac{\left(n+t+k_{1,c-t:n+c}^2\right)}{\left(n+t-1+k_{1,c-t+1:n+c}^2\right)} \left(1-\frac{p}{n+t}\right)\left(1-\frac{h}{n+t+1}\right)\left(1+\frac{p-1}{n+t-p}\right)$$

$$\approx \frac{\left(n+t+k_{1,c-t:n+c}^2\right)}{\left(n+t-1+k_{1,c-t+1:n+c}^2\right)} \left(1-\frac{p}{n+t}-\frac{h}{n+t+1}+\frac{p-1}{n+t-p}\right)$$

$$\approx \frac{\left(n+t+k_{1,c-t:n+c}^2\right)}{\left(n+t-1+k_{1,c-t+1:n+c}^2\right)} \left(1-\frac{1+h}{n+t+1}\right)$$

Let $f(t) > 1$ we get that

$$\frac{\left(n+t+k_{1,c-t:n+c}^2\right)}{\left(n+t-1+k_{1,c-t+1:n+c}^2\right)}\left(1-\frac{1+h}{n+t+1}\right)>1$$

$$k_{1,c-t:n+c}^2(n+t-h)-k_{1,c-t+1:n+c}^2(n+t+1)>h(n+t)-1$$

It can be approximated that

$$k_{1,c-t:n+c}^2-k_{1,c-t+1:n+c}^2>h$$

$$k_{1,c-t:n+c}^2-k_{1,c-t+1:n+c}^2\approx\frac{D_s^2}{c^2(n+t)^2}\ \ [-cn+(c-t-1)n]^2\approx\frac{(t+1)^2n^2D_s^2}{c^2(n+t)^2}>h$$

$$t>t^*=\frac{\frac{\sqrt{h}}{r_s}-1}{1-\frac{\sqrt{h}}{nr_s}}$$

where

$$r_s=\frac{D_s}{c}=\frac{\sqrt{(\mu_0-\mu_b)^T\Sigma^{-1}(\mu_0-\mu_b)}}{c}$$

5. Proof of Theorem 5

Thm 5.1. Suppose $A=P\cdot\mathrm{diag}\left(\alpha_1,\dots,\alpha_p\right)\cdot P^{-1}$. Since $b_t=\mu_t-\mu=A^t(\mu_0-\mu)=A^tb_0=P\cdot\mathrm{diag}\left(\alpha_1^t,\dots,\alpha_p^t\right)P^{-1}b_0$, each component of b_t can be expressed in the form of

$$b_{tj}=\sum_{d=1}^{p}B_{jd}\alpha_d^t$$

where B_{jd} are the constants determined by P and b_0. Since $Y_t-\mu=A^t(Y_0-\mu)+\sum_{l=0}^{t-1}A^lu_{t-l}$, the covariance of X_i and X_j can be derived as

$$\mathrm{cov}\left(X_i,X_j\right)=A^i\Sigma\left(A^j\right)^T+\sum_{s=0}^{\min(i,j)-1}A^s\Sigma_u\left(A^{|i-j|+s}\right)^T \qquad \text{(S5.1)}$$

Based on the equation $\Sigma=A\Sigma A^T+\Sigma_u$, we can get

$$A^i\Sigma\left(A^j\right)^T = \sum_{s=1}^{\infty} A^{i+s}\Sigma_{\mathrm{u}}\left(A^{j+s}\right)^T \tag{S5.2}$$

Using Eq. (S5.1) and (S5.2) we can obtain $\mathrm{cov}\left(X_i, X_j\right) = (c_{ijlm}^2)$ with each entry c_{ijlm} expressed as

$$c_{ijlm}^2 = \sum_{d=1}^{p} C_{lmd}\alpha_d^{|i-j|}, \tag{S5.3}$$

where C_{lmd} are the constants determined by P and Σ_u. Note that if $i = j$, $\mathrm{cov}\left(X_i, X_j\right) = \Sigma = (\sigma_{lm}^2)$. The sample covariance matrix $S_{k+1:n} = \frac{1}{n-k-1}\sum_{i=k+1}^{n}\left(X_i - \bar{X}_{n,k}\right)\left(X_i - \bar{X}_{n,k}\right)^T$. Let

$$S = \frac{n-k-1}{n-k}S_{k+1:n} = \frac{1}{n-k}\sum_{i=k+1}^{n}\left(X_i - \bar{X}_{n,k}\right)\left(X_i - \bar{X}_{n,k}\right)^T = (s_{ij})$$

Denote $y_{ij} = \left(X_i - \bar{X}_{n,k}\right)_j$, i.e., the j-th dimension of $X_i - \bar{X}_{n,k}$, then

$$\beta_{ij} = E\left(y_{ij}\right) = \mu_{ij} - \frac{\sum_{t=k+1}^{n}\mu_{it}}{n-k} = b_{ij} - \frac{\sum_{t=k+1}^{n} b_{tj}}{n-k} = b_{ij} + O(\frac{1}{n-k}) \tag{S5.4}$$

The variance of y_{ij} can be calculated as

$$\gamma_{iijj}^2 = \mathrm{cov}\left(x_{ij} - \bar{x}_j, x_{ij} - \bar{x}_j\right)$$
$$= \sigma_{jj}^2 - \frac{2}{n-k}\sum_{i'=k+1}^{n} c_{i'ijj}^2 + \frac{1}{(n-k)^2}\sum_{i'=k+1}^{n}\sum_{j'=k+1}^{n} c_{i'j'jj}^2 = \sigma_{jj}^2 + O(\frac{1}{n-k}) \tag{S5.5}$$

The cross covariance can be obtained using the same way

$$\gamma_{ii'jj'}^2 = \mathrm{cov}\left(y_{ij}, y_{i'j'}\right) = c_{ii'jj'}^2 + O(\frac{1}{n-k}) \tag{S5.6}$$

To facilitate understanding, let's first consider a relatively simple case with $p = 2$. The case $p = 1$ is much simpler and is not shown here.

$$|S| = \frac{1}{(n-k)^2}\left[\left(\sum_{i=k+1}^{n} y_{i1}^2\right)\left(\sum_{i=k+1}^{n} y_{i2}^2\right) - \left(\sum_{i=k+1}^{n} y_{i1}y_{i2}\right)^2\right] \tag{S5.7}$$

$$= \frac{1}{(n-k)^2}\left[\sum_{i=k+1}^{n}\sum_{j=k+1}^{n} y_{i1}^2 y_{j2}^2 - y_{i1}y_{i2}y_{j1}y_{j2}\right]$$

According to Lemma 1,

$$E\left(y_{i1}^2 y_{j2}^2\right) = \beta_{i1}^2\beta_{j2}^2 + \beta_{i1}^2\gamma_{jj22}^2 + \beta_{j2}^2\gamma_{ii11}^2 + 4\beta_{i1}\beta_{j2}\gamma_{ij12}^2 + \gamma_{ii11}^2\gamma_{jj22}^2 + 2\gamma_{ij12}^4$$

$$E\left(y_{i1}y_{i2}y_{j1}y_{j2}\right)$$
$$= \beta_{i1}\beta_{i2}\beta_{j1}\beta_{j2} + \beta_{i1}\beta_{i2}\gamma_{jj12}^2 + \beta_{i1}\beta_{j1}\gamma_{ij22}^2 + \beta_{i1}\beta_{j2}\gamma_{ij21}^2 + \beta_{i2}\beta_{j1}\gamma_{ij12}^2 + \beta_{i2}\beta_{j2}\gamma_{ij11}^2$$
$$+ \beta_{j1}\beta_{j2}\gamma_{ii12}^2 + \gamma_{ii12}^2\gamma_{jj12}^2 + \gamma_{ij11}^2\gamma_{ij22}^2 + \gamma_{ij12}^2\gamma_{ij21}^2$$

First look at the terms that lead to the order of $O(1)$ after summation and normalization. These terms should be $\sim O(1)$. Therefore they should not contain any β and cross-covariance. Here only $\gamma_{ii11}^2\gamma_{jj22}^2 - \gamma_{ii12}^2\gamma_{jj12}^2$ satisfies the requirements. Besides, the $O(\frac{1}{n-k})$ in Eq. (S5.5) and (S5.6) can all be ignored when calculating these two terms. If we check these terms we will find that we only need to replace the terms in Eq. (S5.7) with the corresponding variance and covariance, e.g., replacing y_{i1}^2 with σ_{11}^2, and replacing $y_{i1}y_{i2}$ with σ_{12}^2. Clearly, after replacement, we will get exactly $\sigma_{11}^2\sigma_{22}^2 - \sigma_{12}^4 = |\Sigma|$ after summation and normalization.

Now check the terms that lead to the order of $O(\frac{1}{n-k})$. In the previous step, we ignored the terms of order $O(\frac{1}{n-k})$ in the calculation of variances and covariance. If we keep these terms, we will obtain a term $\frac{C_1}{n-k}$ where C_1 is a constant independent of k (ignoring the terms α_d^{n-k} as they have much higher shrinking order). Among the remaining terms not used at the previous step, those with two $\beta's$ but without cross-covariance and those only have cross-covariance could achieve the order of $O(\frac{1}{n-k})$. These terms are $\left(\beta_{i1}^2\gamma_{jj22}^2 + \beta_{j2}^2\gamma_{ii11}^2 - \beta_{i1}\beta_{i2}\gamma_{jj12}^2 - \beta_{j1}\beta_{j2}\gamma_{ii12}^2\right) + \left(\gamma_{ij11}^2\gamma_{ij22}^2 + \gamma_{ij12}^2\gamma_{ij21}^2\right)$. In calculating $\left(\beta_{i1}^2\gamma_{jj22}^2 + \beta_{j2}^2\gamma_{ii11}^2 - \beta_{i1}\beta_{i2}\gamma_{jj12}^2 - \beta_{j1}\beta_{j2}\gamma_{ii12}^2\right)$ to achieve the order of $O(\frac{1}{n-k})$, the terms of order $O(\frac{1}{n-k})$ in Eq. (S5.5) and (S5.6) need to be removed. Therefore,

$$\sum_{i=k+1}^{n}\sum_{j=k+1}^{n}\beta_{i1}^2\gamma_{jj22}^2+\beta_{j2}^2\gamma_{ii11}^2-\beta_{i1}\beta_{i2}\gamma_{jj12}^2-\beta_{j1}\beta_{j2}\gamma_{ii12}^2$$
$$=(n-k)\sum_{i=k+1}^{n}\beta_{i1}^2\sigma_{22}^2+\beta_{i2}^2\sigma_{11}^2-\beta_{i1}\beta_{i2}\sigma_{12}^2-\beta_{i1}\beta_{i2}\sigma_{12}^2$$
$$=(n-k)|\Sigma|b_i^T\Sigma^{-1}b_i$$

For the terms with only covariance,

$$\frac{1}{(n-k)^2}\left[\sum_{i=k+1}^{n}\sum_{j=k+1}^{n}\gamma_{ij11}^2\gamma_{ij22}^2+\gamma_{ij12}^2\gamma_{ij21}^2\right]=\frac{C_2}{n-k}$$

where C_2 is a constant independent of k in the order of $O(\frac{1}{n-k})$.

For the cases $p>2$, we can follow exactly the same procedure and get the same result.

$$|S|=\frac{1}{(n-k)^p}\sum_{j_1,\ldots j_p}(-1)^{\tau(j_1,\ldots,j_p)}s_{1j_1}s_{2j_2}\ldots s_{pj_p}$$

where $s_{dj_p}=\sum_{i=k+1}^{n}y_{id}y_{ij_p}$, and $\tau(j_1,\ldots,j_p)$ is the number of inversions of the permutation $(j_1\ldots j_p)$. It can be further written as

$$|S|=\frac{1}{(n-k)^p}\sum_{j_1,\ldots j_p}(-1)^{\tau(j_1,\ldots,j_p)}\sum_{i_1=k+1}^{n}\cdots\sum_{i_p=k+1}^{n}y_{i_1 1}y_{i_1 j_1}y_{i_2 2}y_{i_2 j_2}\cdots y_{i_p p}y_{i_p j_p}$$

Similarly, to get the term of order $O(1)$, the terms in $E(y_{i_1 1}y_{i_1 j_1}y_{i_2 2}y_{i_2 j_2}\cdots y_{i_p p}y_{i_p j_p})$ should all be variances and covariances at the same time step (no cross-covariance). Therefore we get the term $\prod_{d=1}^{p}\sigma_{dj_d}^2$, which eventually leads to $|\Sigma|$. For the term of $|S|$ with order $O(\frac{1}{n-k})$, one part with value $\frac{C_1}{n-k}$ is from $\sum c_{i_1 i_1 1 j_1}^2\cdots c_{i_p i_p p j_p}$ when we keep $O(\frac{1}{n-k})$ in Eq. (S5.5) and (S5.6); the second part with vaule $\frac{C_2}{n-k}$ is from the terms with each consisting of two cross-covariances and $p-2$ variances or covariances without time lag; the third part is from the terms with two $\beta's$ at the same time step and with all others being the variances and covariances at the same time step, e.g., $\beta_{i_1 1}\beta_{i_1 j_1}\sigma_{2j_2}^2\cdots\sigma_{pj_p}^2$ It would be easy to verify that the summation of all terms containing $\beta_{i_s d}\beta_{i_s j_d}$ would be exactly $\beta_{i_s d}\beta_{i_s j_d} * M_{d,j_d}$ where M_{d,j_d} is the minor of the entry in the d-th row and j_d-th column of the matrix Σ. Therefore we can get $\frac{1}{n-k}|\Sigma|\sum_{i=k+1}^{n}b_i^T\Sigma^{-1}b_i$. In summary, the following equation is proved:

$$E|S| = E\left(\left|\frac{n-k-1}{n-k}S_{k+1:n}\right|\right) = |\Sigma|\left[1+\frac{1}{n-k}\sum_{i=k+1}^{n} b_i^T\Sigma^{-1}b_i + \frac{C}{n-k} + O\left(\frac{1}{(n-k)^2}\right)\right]$$

Thm 5.2.

$$R_{k+1:n} = \frac{|S_{k+1:n}|}{(n-k)^h}$$

$$\frac{\mathrm{E}(R_{k+1:n})}{E(R_{k:n})} = \frac{(n-k)^p}{(n-k-1)^p}\frac{(n-k+1)^h}{(n-k)^h}\frac{1+\frac{1}{n-k}\sum_{i=k+1}^{n} b_i^T\Sigma^{-1}b_i + \frac{C}{n-k}}{1+\frac{1}{n-k+1}\sum_{i=k}^{n} b_i^T\Sigma^{-1}b_i + \frac{C}{n-k+1}} \le 1$$

$$b_k^T\Sigma^{-1}b_k \ge h$$

6. Proof of Proposition 1

$$R_{c+1:c+n} = \frac{|S_{c+1:c+n}|}{n^h}, R_{1:c+n} = \frac{|S_{1:c+n}|}{(n+c)^h}$$

Let

$$\bar{X} = \frac{1}{n+c}\sum_{i=1}^{n+c} X_i\,, \qquad \bar{X}_1 = \frac{1}{c}\sum_{i=1}^{c} X_i\,, \qquad \bar{X}_2 = \frac{1}{n}\sum_{i=c+1}^{c+n} X_i$$

Then

$$\bar{X} = \frac{c\bar{X}_1 + n\bar{X}_2}{c+n}$$

$$\begin{aligned}
S_{1:c+n} &= \frac{1}{n+c-1}\sum_{i=1}^{c+n}(X_i-\bar{X})(X_i-\bar{X})^T \\
&= \frac{1}{n+c-1}\sum_{i=1}^{c}(X_i-\bar{X}_1)(X_i-\bar{X}_1)^T + \frac{1}{n+c-1}\sum_{i=c+1}^{c+n}(X_i-\bar{X}_2)(X_i-\bar{X}_2)^T \\
&\quad + \frac{nc}{(n+c-1)(n+c)}(\bar{X}_1-\bar{X}_2)(\bar{X}_1-\bar{X}_2)^T \\
&= \frac{c}{n+c-1}S_{1:c} + \frac{n}{n+c-1}S_{c+1:c+n} + \frac{nc}{(n+c-1)(n+c)}(\bar{X}_1-\bar{X}_2)(\bar{X}_1-\bar{X}_2)^T
\end{aligned}$$

Since $S_{1:c}, S_{c+1:c+n}$ and $(\bar{X}_1-\bar{X}_2)(\bar{X}_1-\bar{X}_2)^T$ are all positive semi-definite, based on the Minkowski inequality, we have

$$|S_{1:c+n}| \geq \left|\frac{c}{n+c-1}S_{1:c}\right| + \left|\frac{n}{n+c-1}S_{c+1:c+n}\right| + \left|\frac{nc}{(n+c-1)(n+c)}(\bar{X}_1-\bar{X}_2)(\bar{X}_1-\bar{X}_2)^T\right|$$

Since $(\bar{X}_1-\bar{X}_2)(\bar{X}_1-\bar{X}_2)^T$ is of rank 1, $\left|\frac{nc}{(n+c-1)(n+c)}(\bar{X}_1-\bar{X}_2)(\bar{X}_1-\bar{X}_2)^T\right| = 0$. Therefore

$$|S_{1:c+n}| \geq \left|\frac{c}{n+c-1}S_{1:c}\right| + \left|\frac{n}{n+c-1}S_{c+1:c+n}\right|$$

$S_{1:c}$ and $S_{c+1:c+n}$ follow Wishart distribution of degree $c-1$ and $n-1$, respectively. Therefore

$$E(|S_{1:c}|) = \frac{\prod_{i=1}^{p}(c-i)}{(c-1)^p}|\Sigma_1|, E(|S_{c+1:c+n}|) = \frac{\prod_{i=1}^{p}(n-i)}{(n-1)^p}|\Sigma_0|$$

$$E(R_{1:c+n}) \geq \frac{1}{(n+c)^h}\left(\frac{c}{n+c-1}\right)^p E(|S_{1:c}|) + \frac{1}{(n+c)^h}\left(\frac{n}{n+c-1}\right)^p E(|S_{c+1:c+n}|) = RHS$$

Let $RHS > E(R_{c+1:c+n})$ we can get

$$E(R_{1:c+n}) \geq \frac{1}{(n+c)^h}\left(\frac{c}{n+c-1}\right)^p E(|S_{1:c}|) + \frac{1}{(n+c)^h}\left(\frac{n}{n+c-1}\right)^p E(|S_{c+1:c+n}|) > \frac{E(|S_{c+1:c+n}|)}{n^h},$$

$$\frac{1}{(n+c)^h}\left(\frac{c}{n+c-1}\right)^p E(|S_{1:c}|) > \left(\frac{1}{n^h} - \frac{1}{(n+c)^h}\left(\frac{n}{n+c-1}\right)^p\right)E(|S_{c+1:c+n}|)$$

$$\frac{1}{(n+c)^h}\left(\frac{c}{n+c-1}\right)^p \frac{\prod_{i=1}^{p}(c-i)}{(c-1)^p}|\Sigma_1| > \left(\frac{1}{n^h} - \frac{1}{(n+c)^h}\left(\frac{n}{n+c-1}\right)^p\right)\frac{\prod_{i=1}^{p}(n-i)}{(n-1)^p}|\Sigma_0|$$

$$\frac{|\Sigma_1|}{|\Sigma_0|} > \frac{\left(\frac{1}{n^h} - \frac{1}{(n+c)^h}\left(\frac{n}{n+c-1}\right)^p\right)\frac{\prod_{i=1}^{p}(n-i)}{(n-1)^p}}{\frac{1}{(n+c)^h}\left(\frac{c}{n+c-1}\right)^p \frac{\prod_{i=1}^{p}(c-i)}{(c-1)^p}} = \frac{\left(\left(1+\frac{c}{n}\right)^h - \left(\frac{n}{n+c-1}\right)^p\right)\frac{\prod_{i=1}^{p}(n-i)}{(n-1)^p}}{\left(\frac{c}{n+c-1}\right)^p \frac{\prod_{i=1}^{p}(c-i)}{(c-1)^p}}$$

References

Anderson, Theodore Wilbur, and Etats-Unis Mathématicien. 1958. *An introduction to multivariate statistical analysis*. Vol. 2: Wiley New York.

Anderson, TW. 1946. "The non-central Wishart distribution and certain problems of multivariate statistics." *The Annals of Mathematical Statistics*:409-431.

Izenman, Alan Julian. 2008. *Modern multivariate statistical techniques*. Vol. 1: Springer.