Asymptotic distribution for centered higher-order moments using Cramer's method

1 Notation and asymptotic distribution of p_k

Let x be a p-dimensional vector variable with finite higher order moments $\nu_k = E(x-\mu)^{\otimes k}$ ($a^{\otimes k} = a \otimes \ldots \otimes a$ is a kronecker product with k terms). Note that $\nu_1 = 0$.
Let $x_1, \ldots, x_n$ be an iid sample of x and

$$p_k = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^{\otimes k},\; k \geq 1$$

By the Law of Large numbers,

$$p_k \xrightarrow{P} \nu_k, \tag{1}$$

and by the Central Limit theorem,

$$\sqrt{n}(p_k - \nu_k) \xrightarrow{D} \mathcal{N}(0, \mathrm{cov}((x_i-\mu)^{\otimes k})), \tag{2}$$

since the $(x_i-\mu)^{\otimes k}$ are iid with variance matrix $\mathrm{cov}((x_i-\mu)^{\otimes k})$, assumed to be finite. This implies that the asymptotic variance matrix $\mathrm{acov}(p_k) = \mathrm{cov}((x_i-\mu)^{\otimes k})$ can be estimated consistenly by the sample variance matrix $\mathrm{scov}((x_i-\mu)^{\otimes k})$ of the $(x_i-\mu)^{\otimes k}$'s.
Let us now define $q_1 = \bar{x}$ and

$$q_k = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^{\otimes k},\; k \geq 2 \tag{3}$$

We will show that like p_k, q_k is an asymptotically normal vector with the same asymptotic limit ν_k as p_k (when $k \neq 1$) but with an asymptotic variance matrix that may differ from the one of p_k when $k > 2$. We also derive an expression for a consistent estimator of $\mathrm{acov}(q)$ when q is an stacked vector of q_ks.

2 Consistent estimation of $\Gamma_p = \mathbf{acov}(p)$

Let

$$p = \begin{pmatrix} p_1 \\ \vdots \\ p_m \end{pmatrix}$$

Note that

$$\sqrt{n}\begin{pmatrix} p_1-\nu_1 \\ \vdots \\ p_m-\nu_m \end{pmatrix} = \frac{1}{\sqrt{n}}\sum_i \begin{pmatrix} x_i-\mu-\nu_1 \\ \vdots \\ (x_i-\mu)^{\otimes m}-\nu_m \end{pmatrix}$$

By the central limit theorem the asymptotic covariance matrix for p is

$$\Gamma_p = \text{cov}\begin{pmatrix} x-\mu \\ \vdots \\ (x-\mu)^{\otimes m} \end{pmatrix}$$

This consists of sub-matrics of the form

$$\Gamma_{rs} = \text{scov}((x-\mu)^{\otimes r},(x-\mu)^{\otimes s}) = E((x-\mu)^{\otimes r})((x-\mu)^{\otimes s})' - E((x-\mu)^{\otimes r})(E(x-\mu)^{\otimes s})'$$

It follows that

$$\text{vec}\,(\Gamma_{rs}) = E(x-\mu)^{\otimes(r+s)} - E((x-\mu)^{\otimes r})\otimes E(x-\mu)^{\otimes s} = \nu_{r+s}-\nu_r\otimes\nu_s \overset{a}{=} q_{r+s}-q_r\otimes q_s$$

This provides a consistent estimator for Γ_{rs} and these provide a consistent estimator for Γ. With a bit of work, we can show that this consistent estimator of Γ_p is the sample covariance matrix scov (t_i^+), where

$$t_i^+ = \begin{pmatrix} x_i-\bar{x} \\ (x_i-\bar{x})^{\otimes 2} \\ \vdots \\ (x_i-\bar{x})^{\otimes k} \end{pmatrix}$$

Unfortunately, what we need is not a consistent estimator of Γ_p but a consistent estimator of Γ_q. This estimator is to be developed below.

3 Representation of q_k as a function of p_1 and p_k

Replacing in (3) $x_i-\bar{x}$ by $(x_i-\mu)-(\bar{x}-\mu)$, and expanding the kronecker power[1] yields

$$(x_i-\bar{x})^{\otimes k} = (x_i-\mu)^{\otimes k} - \sum_{j=1}^{k}(x_i-\mu)^{\otimes(j-1)}\otimes(\bar{x}-\mu)\otimes(x_i-\mu)^{\otimes(k-j)} + r_{ki} \quad (4)$$

[1] Note that the binomial theorem can not be used here since kronecker power is not commutative in general; e.g., $(a+b)^{\otimes 2} \neq a^{\otimes 2}+2a\otimes b+b^{\otimes 2}$, for general vectors a and b

where it is understood that $a^{\otimes 0} = 1$. In formula (4), r_{ki} is a kronecker product of k terms which can be shown to have the term $(x_i - \mu)$ repeated twice or more. For developments to follow, we need the commutation matrix K_{mn} (see Magnus and Neudecker, 1986) defined by

$$K_{mn}\text{vec}(A) = \text{vec}(A'),$$

for any $m \times n$ matrix A. It is known that

$$K_{mn} = \sum_{i=1}^{m}\sum_{j=1}^{n}(H_{ij} \otimes H'_{ji})$$

where H_{ij} is an $m \times n$ matrix with a 1 in the ijth position and zeros elsewhere. If a and b are columns vectors of length n and m then

$$K_{mn}(a \otimes b) = b \otimes a \tag{5}$$

The property (5) permits writing (4) as

$$(x_i - \bar{x})^{\otimes k} = (x_i - \mu)^{\otimes k} - C_k(x_i - \mu)^{\otimes(k-1)} \otimes (\bar{x} - \mu) + r_{ki}, \tag{6}$$

where

$$C_k = \sum_{j=1}^{k-1} K_{p^j p^{k-j}} + I_{p^k}\,, k \geq 2 \tag{7}$$

To achieve this result we used that for $j = 1, \ldots, k-1$

$$\begin{aligned}
&(x_i - \mu)^{\otimes(j-1)} \otimes (\bar{x} - \mu) \otimes (x_i - \mu)^{\otimes(k-j)} \\
&= \left[(x_i - \mu)^{\otimes(j-1)} \otimes (\bar{x} - \mu)\right] \otimes (x_i - \mu)^{\otimes(k-j)} \\
&= K_{p^j p^{k-j}}\left((x_i - \mu)^{\otimes(k-j)} \otimes \left[(x_i - \mu)^{\otimes(j-1)} \otimes (\bar{x} - \mu)\right]\right) \\
&= K_{p^j p^{k-j}}\left((x_i - \mu)^{\otimes(k-j)} \otimes (x_i - \mu)^{\otimes(j-1)} \otimes (\bar{x} - \mu)\right) \\
&= K_{p^j p^{k-j}}\left((x_i - \mu)^{\otimes(k-1)} \otimes (\bar{x} - \mu)\right)
\end{aligned}$$

in virtue of the associative property of the Kronecker product and (5). Clearly, for $j = k$

$$\begin{aligned}
&(x_i - \mu)^{\otimes(j-1)} \otimes (\bar{x} - \mu) \otimes (x_i - \mu)^{\otimes(k-j)} \\
&= (x_i - \mu)^{\otimes(k-1)} \otimes (\bar{x} - \mu)
\end{aligned}$$

Averaging (6) across i and using the definition of the p_k's, we obtain

$$q_k = p_k - C_k\,(p_{k-1} \otimes p_1) + r_k \tag{8}$$

where

$$r_k = \frac{1}{n}\sum_{i=1}^{n} r_{ki}$$

Substracting ν_k in both sides of (8) and multiplying by $\sqrt{n}$, yields

$$\sqrt{n}(q_k - \nu_k) = \sqrt{n}(p_k - \nu_k) - C_k\left(p_{k-1} \otimes \sqrt{n}p_1\right) + \sqrt{n}r_k\,, k \geq 2; \tag{9}$$

thus,

$$\sqrt{n}(q_k - \nu_k) = \left(\begin{array}{cc} -C_k(p_{k-1} \otimes I_p) & I_{p^k} \end{array}\right)\sqrt{n}\left(\begin{array}{c} p_1 - \nu_1 \\ p_k - \nu_k \end{array}\right) + \sqrt{n}r_k\,, k \geq 2 \tag{10}$$

since $\nu_1 = 0$ and

$C_k\left(p_{k-1} \otimes \sqrt{n}(p_1 - \nu_1)\right) = C_k\text{vec}\left(\sqrt{n}(p_1 - \nu_1)p'_{k-1}\right) = C_k(p_{k-1} \otimes I_p)\sqrt{n}(p_1 - \nu_1).$

We now need proving that

$$\sqrt{n}r_k \xrightarrow{P} 0,\ k \geq 2; \tag{11}$$

that is, $\sqrt{n}r_k = o_p(1)$. This is equivalent to proving

$$\frac{1}{\sqrt{n}}\sum r_{ki} \xrightarrow{p} 0 \tag{12}$$

For this we need a little more precise definition of r_{ki}. Let r_{ki} be the sum of all the kth order Kronecker products containing terms of the form $(x_i - \mu)$ or $(\bar{x} - \mu)$ with at least two of these being $(\bar{x} - \mu)$. Each of these is a permutation of

$$(\bar{x} - \mu)^{\otimes j} \otimes (x_i - \mu)^{\otimes (k-j)}$$

for some $j \geq 2$. It follows that

$$\frac{1}{\sqrt{n}}\sum_i (\bar{x} - \mu)^{\otimes j} \otimes (x_i - \mu)^{\otimes (k-j)} = \sqrt{n}(\bar{x} - \mu)^{\otimes j} \otimes \frac{1}{n}\sum_i (x_i - \mu)^{\otimes (k-j)} \xrightarrow{P} 0$$

and (11) follows from this.

Combining (1), (8) and (11), we obtain

$$q_k \xrightarrow{P} \nu_k, \quad k > 1 \tag{13}$$

That is, p_k and q_k have the same asymptotic limit when $k > 1$; when $k = 1$, then $q_1 \xrightarrow{P} \mu$ and $p_1 \xrightarrow{P} \nu_1 = 0$.

By setting $k = 2$ in (9) and using (11) and $p_1 \xrightarrow{P} 0$, we obtain

$$\sqrt{n}(q_2 - \nu_2) = \sqrt{n}(p_2 - \nu_2) + o_p(1). \tag{14}$$

When $k = 1$ then, clearly,

$$\sqrt{n}(q_1 - \mu_1) = \sqrt{n}(p_1 - \nu_1) \tag{15}$$

The results (10), (14) and (15) will be exploited in the next section to derive the asymptotic distribution of a stacked vector of q_k's.

4 Consistent estimation of Γ_q

For simplicity of exposition, consider $q = (q_1, q_2, q_3, q_4)'$, i.e. we consider the case of $k = 4$ (larger values for k would be handled by analogy). By stacking equations (15) and (14), and (10) with $k = 3, 4$, and denoting $\nu = (\nu_1, \nu_2, \nu_3, \nu_k)'$ and $\tilde{\nu} = (\mu_1, \nu_2, \nu_3, \nu_k)'$, since $p \xrightarrow{P} \nu_k$ and $q \xrightarrow{P} \tilde{\nu_k}$, we obtain

$$\sqrt{n}\begin{pmatrix} q_1 - \mu_1 \\ q_2 - \nu_2 \\ q_3 - \nu_3 \\ q_4 - \nu_4 \end{pmatrix} = \begin{pmatrix} I_p & 0 & 0 & 0 \\ 0 & I_{p^2} & 0 & 0 \\ -C_k(p_2 \otimes I_p) & 0 & I_{p^3} & 0 \\ -C_k(p_3 \otimes I_p) & 0 & 0 & I_{p^4} \end{pmatrix} \sqrt{n}\begin{pmatrix} p_1 - \nu_1 \\ p_2 - \nu_2 \\ p_3 - \nu_3 \\ p_4 - \nu_4 \end{pmatrix} + o_p(1); \tag{16}$$

which, by direct application of Slutsky's Theorem, proves that $\sqrt{n}(q - \tilde{\nu})$ is asymptotically normal with variance matrix

$$\Gamma_q = U\Gamma_p U', \tag{17}$$

where

$$U = \begin{pmatrix} I_p & 0 & 0 & 0 \\ 0 & I_{p^2} & 0 & 0 \\ -C_k(\nu_2 \otimes I_p) & 0 & I_{p^3} & 0 \\ -C_k(\nu_3 \otimes I_p) & 0 & 0 & I_{p^4} \end{pmatrix}$$

Since

$$\hat{U} = \begin{pmatrix} I_p & 0 & 0 & 0 \\ 0 & I_{p^2} & 0 & 0 \\ -C_k(q_2 \otimes I_p) & 0 & I_{p^3} & 0 \\ -C_k(q_3 \otimes I_p) & 0 & 0 & I_{p^4} \end{pmatrix}$$

is a consistent estimator of U (we used (13)), a consistent estimate of Γ_q is

$$\hat{\Gamma}_q = \hat{U}[\text{scov}(t_i^+)]\hat{U}' = \text{scov}(\hat{U}t_i^+)$$

That is, $\hat{\Gamma}_q$ can be written as the sample variance matrix of the pseudo-values

$$\tilde{t}_i^+ = \begin{pmatrix} I_p & 0 & 0 & 0 \\ 0 & I_{p^2} & 0 & 0 \\ -C_k(q_2 \otimes I_p) & 0 & I_{p^3} & 0 \\ -C_k(q_3 \otimes I_p) & 0 & 0 & I_{p^4} \end{pmatrix} \begin{pmatrix} x_i - \bar{x} \\ (x_i - \bar{x})^{\otimes 2} \\ (x_i - \bar{x})^{\otimes 3} \\ (x_i - \bar{x})^{\otimes 4} \end{pmatrix} \tag{18}$$

With a bit of work it can be seen that $\text{scov}(\tilde{t}_i^+)$ coincides with the accelerated version of the IJK estimator of variance proposed in our paper, the "accelerating" matrix C_k being now expressed in terms of the commutation matrices K_{mn}. Using the K_{mn}'s can be less efficient (computationally) than using the matrix C_k defined in the Appendix of the paper.