A detailed derivation of $\hat{q}_n \to \tilde{q}_n$ in probability

The derivation of the fact that $\hat{q}_n \to \tilde{q}_n$ in probability can be established in a similar way as in that of Theorem 2.1 of Srivastava and Du (2008). Since $\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2 \sim N_p(\boldsymbol{0}, N_1^{-1}\boldsymbol{\Sigma}_1 + N_2^{-1}\boldsymbol{\Sigma}_2)$ we have $\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2 = \boldsymbol{D}^{1/2}\boldsymbol{y}$ where $\boldsymbol{y} \sim N_p(\boldsymbol{0}, \boldsymbol{\mathcal{R}})$. Then $(\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)^T \hat{\boldsymbol{D}}^{-1}(\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2) = \boldsymbol{y}^T \hat{\boldsymbol{D}}^{-1}\boldsymbol{D}\boldsymbol{y}$. Let $\sigma_{ii} = N_1^{-1}\sigma_{1ii} + N_2^{-1}\sigma_{2ii}$ and $s_{ii} = N_1^{-1}s_{1ii} + N_2^{-1}s_{2ii}$. Then $\hat{\boldsymbol{D}}\boldsymbol{D}^{-1} = \mathrm{diag}(\sigma_{11}/s_{11}, \ldots, \sigma_{pp}/s_{pp})$. Thus, by noting that

$$\frac{\sigma_{ii}}{s_{ii}} = 1 - \left(\frac{s_{ii}}{\sigma_{ii}} - 1\right) + \left(\frac{\sigma_{ii}}{s_{ii}} + \frac{s_{ii}}{\sigma_{ii}} - 2\right),$$

we have

$$\begin{aligned}
\hat{q}_n &= \frac{(\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)^T \hat{\boldsymbol{D}}^{-1}(\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2) - p}{\sqrt{p}} = \frac{\boldsymbol{y}^T \hat{\boldsymbol{D}}^{-1}\boldsymbol{D}\boldsymbol{y} - p}{\sqrt{p}} \\
&= \frac{\boldsymbol{y}^T\boldsymbol{y} - p}{\sqrt{p}} - \frac{\boldsymbol{y}^T\boldsymbol{D}_1\boldsymbol{y}}{\sqrt{p}} + \frac{\boldsymbol{y}^T\boldsymbol{D}_2\boldsymbol{y}}{\sqrt{p}} \\
&= A - B + C, \quad say,
\end{aligned}$$

where

$$\boldsymbol{D}_1 = \mathrm{diag}\left(\frac{s_{11}}{\sigma_{11}} - 1, \ldots, \frac{s_{pp}}{\sigma_{pp}} - 1\right), \quad \boldsymbol{D}_2 = \mathrm{diag}\left(\frac{\sigma_{11}}{s_{11}} + \frac{s_{11}}{\sigma_{11}} - 2, \ldots, \frac{\sigma_{pp}}{s_{pp}} + \frac{s_{pp}}{\sigma_{pp}} - 2\right).$$

Note that $A = \tilde{q}_n$. Let $u_i = s_{ii}/\sigma_{ii} - 1$, then $E(u_i) = 0$ and $E(u_i u_j) \leq 2\rho_{ij}^2/\min(n_1, n_2)$. Since $\boldsymbol{y}$ is independent of $\boldsymbol{D}_1$ and $\boldsymbol{y} = \boldsymbol{\mathcal{R}}^{1/2}\boldsymbol{z}$ where $\boldsymbol{z} \sim N_p(\boldsymbol{0}, \boldsymbol{I}_p)$, we have $E(\boldsymbol{y}^T\boldsymbol{D}_1\boldsymbol{y}) = E\{\boldsymbol{y}^T E(\boldsymbol{D}_1)\boldsymbol{y}|\boldsymbol{y}\} = 0$ and

$$\begin{aligned}
\mathrm{Var}(\boldsymbol{y}^T\boldsymbol{D}_1\boldsymbol{y}) &= E\{(\boldsymbol{z}^T\boldsymbol{\mathcal{R}}^{1/2}\boldsymbol{D}_1\boldsymbol{\mathcal{R}}^{1/2}\boldsymbol{z})^2\} = E\{2\mathrm{tr}(\boldsymbol{\mathcal{R}}\boldsymbol{D}_1)^2 + (\mathrm{tr}\boldsymbol{\mathcal{R}}\boldsymbol{D}_1)^2\} \\
&= E\left\{2\sum_{i,j=1}^{p}\rho_{ij}^2 u_i u_j + \left(\sum_{i=1}^{p} u_i\right)^2\right\} = E\left\{3\sum_{i=1} u_i^2 + \sum_{i\neq j}(1 + 2\rho_{ij}^2)u_i u_j\right\} \\
&\leq \frac{6p}{\min(n_1, n_2)} + \frac{2}{\min(n_1, n_2)}\sum_{i\neq j}(1 + 2\rho_{ij}^2)\rho_{ij}^2 \\
&\leq \frac{3}{\min(n_1, n_2)}\left\{2p + \sum_{i\neq j}\rho_{ij}^2 + \sum_{i\neq j}\rho_{ij}^2\rho_{ij}^2\right\} \\
&\leq \frac{6}{\min(n_1, n_2)}\mathrm{tr}\boldsymbol{\mathcal{R}}^2.
\end{aligned}$$

In the second equality above we use the fact that

$$E\{(\boldsymbol{z}^T\boldsymbol{A}\boldsymbol{z})^2\} = E\{2\mathrm{tr}\boldsymbol{A}^2 + (\mathrm{tr}\boldsymbol{A})^2\}$$

for a symmetric matrix $\boldsymbol{A}$ independent of $\boldsymbol{z}$ (see, e.g., Schott (2005)) and we use $\rho_{ij}^2 \leq 1$ in the last inequality. Thus, we obtain $B = O_p(N_m^{-1/2})$ if $\mathrm{tr}\boldsymbol{\mathcal{R}}^2 = O(p)$ where $N_m = \min(N_1, N_2)$.

Since $\boldsymbol{D}_2$ is non-negative matrix, it follows that

$$E(|\boldsymbol{y}^T\boldsymbol{D}_2\boldsymbol{y}|) = E(\boldsymbol{y}^T\boldsymbol{D}_2\boldsymbol{y}) = E(\mathrm{tr}\boldsymbol{\mathcal{R}}\boldsymbol{D}_2) = \sum_{i=1}^{p} E\left(\frac{\sigma_{ii}}{s_{ii}} + \frac{s_{ii}}{\sigma_{ii}} - 2\right).$$

Note that $E(s_{ii}/\sigma_{ii} - 1) = 0$. By the taylor expansion we also note that $E(\sigma_{ii}/s_{ii}) = 1 + O(N_m^{-1})$. Thus, $E(|\boldsymbol{y}^T\boldsymbol{D}_2\boldsymbol{y}|) = O(pN_m^{-1})$ which implies $C = O_p(p^{1/2}N_m^{-1})$ form the Markov inequality. Therefore,

$$\hat{q}_n = \tilde{q}_n + O_p\left\{\max\left(\frac{1}{\sqrt{N_m}}, \frac{\sqrt{p}}{N_m}\right)\right\}.$$

References

J.R. Schott. *Matrix analysis for statistics, 2nd edition*. New York: Wiley, 2005.

M.S. Srivastava and M. Du. A test for the mean vector with fewer observations than the dimension. *Journal of Multivariate Analysis*, 99(3):386–402, 2008.