Testing block independence, $K$ blocks.

Consider the MANOVA model given by the subspace $L \subset \mathbb{R}^N$ ($l := \text{Dim}(L)$) and the finite set $I$ (the index set for the variates), i.e., the model

$$(N_{\xi, \text{Diag}(\Sigma|\nu \in N)} \in \mathcal{P}(\mathbb{R}^{N \times I})| (\xi, \Sigma) \in L^I \times \mathcal{P}(I)).$$

We already know the complete solution to this model: If $N \geq I - l$ the ML estimator $(\hat{\xi}(x), \hat{\Sigma}(x))$ for $(\xi, \Sigma) \in L^I \times \mathcal{P}(I)$ exists with probability one (in the opposite case $N < I - l$ it does not exists for any observation $x \in \mathbb{R}^{N \times I}$) and is uniquely given by

$$\hat{\xi}(x) = Px \quad \text{and} \quad \hat{\Sigma}(x) = \frac{1}{N}(x^tx - x^tPx).$$

The maximum value of the Likelihood function is:

$$L((\hat{\xi}(x), \hat{\Sigma}(x)), x) = \frac{1}{\left(\sqrt{2\pi}\right)^{NI} |\hat{\Sigma}(x)|^{N/2}} \cdot \exp\left\{-\frac{NI}{2}\right\}.$$
Let $K$ be a finite set with $K \geq 2$. Suppose that $I$ is decomposed into a disjoint union of non-empty sets indexed by $k \in K$, i.e., $I = \dot{\bigcup}(I_k|k \in K)$. The parameter $\Sigma \in \mathcal{P}(I)$ can then be partitioned into $K \times K$ block matrix:

$$\Sigma \equiv (\Sigma_{kk'}|(k, k') \in K \times K),$$

where for $k, k' \in K$, $\Sigma_{kk'} \in \mathbb{R}^{I_k \times I_{k'}}$ with $\Sigma_{kk'} = \Sigma_{k'k}$, and $\Sigma_k := \Sigma_{kk'} \in \mathcal{P}(I_k)$.

We will now test the hypothesis $H_0$: $\Sigma_{kk'} = 0$, $k, k' \in K$ with $k \neq k'$. Thus the $I \times I$ variance matrix $\Sigma$ is assumed to have the form:

$$\Sigma \equiv \text{Diag}(\Sigma_k|k \in K).$$

This means that the observables $x_{\nu i_1} \in \mathbb{R}$ and $x_{\nu i_2} \in \mathbb{R}$, $x \equiv (x_{\nu i}|(\nu, i) \in N \times I)$ are independent when $i_1 \in I_k$ and $i_{k'} \in I_{k'}$ with $k \neq k'$, $k, k' \in K$. Note that $L^I = L^{\dot{\bigcup}(I_k|k \in K)} = \times(L^I_k|k \in K)$ and that $\mathbb{R}^{N \times I} = \times(\mathbb{R}^{N \times I_k}|k \in K)$. Thus $\xi = (\xi_k|k \in K)$ with $\xi_k \in L^I_k$, for $\xi \in L^I$ and $x = (x_k|k \in K)$ with $x_k \in \mathbb{R}^{N \times I_k}$, for $x \in \mathbb{R}^{N \times I}$. 
Testing block independence, $K$ blocks.(cont)

Under $H_0$ we then have the model

$$(\otimes (N_{\xi_k}, \text{Diag}(\Sigma_k | \nu \in N) | k \in K) \in \mathcal{P}(\times (\mathbb{R}^{N \times I_k} | k \in K)))|$$

$$((\xi_k | k \in K), (\Sigma_k | k \in K)) \in (\times (L^{I_k} | k \in K)) \times (\times (P(I_k) | k \in K))).$$

This is (formally) NOT a MANOVA model, but it is a product\(^1\) of $K$ MANOVA models indexed by $k \in K$, the model corresponding to $k \in K$ is given by $L \subset \mathbb{R}^N$ and $I_k$. In other words the likelihood function for the model $H_0$ is

$$L((\xi_k | k \in K), (\Sigma_k | k \in K), (x_k | k \in K)) = \prod (L_k((\xi_k, \Sigma_k), x_k) | k \in K),$$

\(\text{where } L_k, k \in K\) are the likelihood functions for the $K$ MANOVA problems. Finding the maximizer for $L(\bullet, (x_k | k \in K))$ then reduces to the wellknown problem of finding maximizers for $L_k(\bullet, x_k), k \in K$.

Note that the condition, $N \geq I - l$, for the existence of the ML estimator under $H$ implies the existence $(N \geq I_k - l, k \in K)$ of the maximizers for $L_k(\bullet, x_k), k \in K$.

\(^1\)We have not formally defined a product of $K$ statistical models and we shall not do so.
Thus under $H_0$ we have

$$\hat{\xi}_k((x_{k'}|k' \in K)) = P x_k, \quad \hat{\Sigma}_k((x_{k'}|k' \in K)) = \frac{1}{N}(x_k^t x_k - x_k^t P x_k), \quad k \in K.$$ 

Note that $\hat{\Sigma}_k = (\hat{\Sigma})_{kk}, \quad k \in K$.

The maximum value of the Likelihood function at $x \equiv (x_k|k \in K) \in \times (\mathbb{R}^{N \times I_k}|k \in K)$ is

$$L(((\hat{\xi}_k(x)|k \in K), (\hat{\Sigma}_k(x)|k \in K)), x) = \prod_{k}(L_k((\hat{\xi}_k(x_k), \hat{\Sigma}_k(x_k)), x_k)|k \in K) = \prod_{k}
\left(\frac{1}{\sqrt{2\pi}^{NI_k}} \frac{1}{|\hat{\Sigma}_k(x_k)|^{N/2}} \exp\left\{-\frac{NI_k}{2}\right\}|k \in K\right) = \frac{1}{\sqrt{2\pi}^{NI} \prod(|\hat{\Sigma}_k(x_k)|^{N/2}|k \in K)} \exp\left\{-\frac{NI}{2}\right\}. \right.$$
Testing block independence, $K$ blocks. The LR statistics.

The LR test statistics at the observation point $x \equiv (x_k | k \in K) \in \times (\mathbb{R}^{N \times I_k} | k \in K)$ is thus

$$q(x) = \frac{1}{\sqrt{2\pi}^{NI}} \frac{\prod_{k \in K} |\hat{\Sigma}_k(x_k)|^{N/2}}{\prod_{k \in K} |\hat{\Sigma}(x)|^{N/2}} \exp\{-\frac{NI}{2}\} \exp\{-\frac{NI}{2}\} =$$

$$\left(\frac{|\hat{\Sigma}(x)|}{\prod_{k \in K} |\hat{\Sigma}_k(x_k)|^{N/2}}\right)^{N/2}.$$  

The equivalent test statistics $-2 \log(q)$ with large values extreme is thus

$$-2 \log(q(x)) = N\left(\sum (\log(|\hat{\Sigma}_k(x_k)|) | k \in K) - \log(|\hat{\Sigma}(x)|)\right).$$

The approximative distribution of this test statistic is a Chi-square distribution with the number of free parameters under H, minus the number of free parameters under $H_0$, i.e.,

$$f := [Il + I(I + 1)/2] - \left[\sum (I_k l | k \in K) + \sum (I_k (I_k + 1)/2 | k \in K)\right] = \frac{1}{2} \sum (I_k I_{k'} | k, k' \in K, k \neq k')$$ as degrees of freedom.
Testing identity of two variance matrices (Bartlett’s test)

Consider two MANOVA models, one given by the subspace $L_1 \subset \mathbb{R}^{N_1}$ and the set $I$, and one given by the subspace $L_2 \subset \mathbb{R}^{N_2}$ and the SAME set $I$. Set $l_1 := \text{Dim}(L_1)$ and $l_2 := \text{Dim}(L_2)$. The product of these two MANOVA models

$$(N_{\xi_1, \text{Diag}(\Sigma_1|\nu_1 \in N_1)} \in \mathcal{P}(\mathbb{R}^{N_1 \times I})|((\xi_1, \Sigma_1) \in L_1^I \times \mathcal{P}(I))$$

and

$$(N_{\xi_2, \text{Diag}(\Sigma_2|\nu_1 \in N_2)} \in \mathcal{P}(\mathbb{R}^{N_2 \times I})|((\xi_2, \Sigma_2) \in L_2^I \times \mathcal{P}(I))$$

expresses that the two corresponding data sets $x_1 \in \mathbb{R}^{N_1 \times I}$ and $x_2 \in \mathbb{R}^{N_2 \times I}$ can be considered as result independent random points in $\mathbb{R}^{N_1 \times I}$ and $\mathbb{R}^{N_2 \times I}$, respectively. More precisely we consider the model (hypothesis H the big model):

$$(N_{\xi_1, \text{Diag}(\Sigma_1|\nu_1 \in N_1)} \otimes N_{\xi_2, \text{Diag}(\Sigma_2|\nu_2 \in N_2)} \in \mathcal{P}(\mathbb{R}^{N_1 \times I} \times \mathbb{R}^{N_2 \times I})|((\xi_1, \xi_2, \Sigma_1, \Sigma_2) \in L_1^I \times L_2^I \times \mathcal{P}(I) \times \mathcal{P}(I)),$$
As before the likelihood function of the last model is an independent product of the likelihood functions of the two MANOVA models. The condition(s) for the existence of the ML estimator is $N_1 \geq I + l_1$ and $N_2 \geq I + l_2$ and the ML estimator at

$$x \equiv (x_1^t, x_2^t)^t = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^{N_1 \times I} \times \mathbb{R}^{N_2 \times I} \equiv \mathbb{R}^{(N_1 \cup N_2) \times I} \equiv \mathbb{R}^{N \times I}$$

with $N := N_1 \cup N_2$, of the last model is thus obtained by the similar quantities of the two MANOVA models, i.e.,

$$\hat{\xi}_1(x) = P_1 x_1, \quad \hat{\Sigma}_1(x) = \frac{1}{N_1} (x_1^t x_1 - x_1^t P_1 x_1),$$

$$\hat{\xi}_2(x) = P_2 x_2, \quad \hat{\Sigma}_2(x) = \frac{1}{N_2} (x_2^t x_2 - x_2^t P_2 x_2),$$

where $P_1$ and $P_2$ are the orthogonal $N_1 \times N_1$ and $N_2 \times N_2$ matrices corresponding to $L_1$ and $L_2$, respectively. The maximum value becomes:

$$\frac{1}{\sqrt{2\pi}^{N_1 I/2}} \frac{1}{\left| \hat{\Sigma}_1(x_1) \right|^{N_1/2}} \exp\left\{-\frac{N_1 I}{2} \right\} \cdot \frac{1}{\sqrt{2\pi}^{N_2 I/2}} \frac{1}{\left| \hat{\Sigma}_2(x_2) \right|^{N_2/2}} \exp\left\{-\frac{N_2 I}{2} \right\} = \frac{1}{\sqrt{2\pi}^{NI/2}} \frac{1}{\left| \hat{\Sigma}_1(x_1) \right|^{N_1/2} \left| \hat{\Sigma}_2(x_2) \right|^{N_2/2}} \exp\left\{-\frac{NI}{2} \right\}. $$
Now consider the hypothesis (identity of the variance matrices in the two MANOVA models)

\[ H_0 : \Sigma_1 = \Sigma_2 =: \Sigma \]

i.e., the small model

\[ \left( N_{\xi_1, \text{Diag}(\Sigma|\nu_1 \in N_1)} \otimes N_{\xi_2, \text{Diag}(\Sigma|\nu_2 \in N_2)} \right) \in \mathcal{P}(\mathbb{R}^{N_1 \times I} \times \mathbb{R}^{N_2 \times I}) | \]

\[ (\xi_1, \xi_2, \Sigma) \in L_1^I \times L_2^I \times \mathcal{P}(I). \]

Now note that

\[ \mathbb{R}^{N_1 \times I} \times \mathbb{R}^{N_2 \times I} = \mathbb{R}^{N \times I}, \]

\[ N_{\xi_1, \text{Diag}(\Sigma|\nu_1 \in N_1)} \otimes N_{\xi_2, \text{Diag}(\Sigma|\nu_2 \in N_2)} = N_{\xi, \text{Diag}(\Sigma|\nu \in N)}, \text{ with } \xi := \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}, \]

and

\[ L_1^I \times L_2^I = (L_1 \times L_2)^I = L^I, \text{ with } L := L_1 \times L_2 \subset \mathbb{R}^{N_1} \times \mathbb{R}^{N_2} \equiv \mathbb{R}^{N}. \]

the small model can be rewritten as the MANOVA model given by \( L \subset \mathbb{R}^{N} \) and \( I \)

\[ (N_{\xi, \text{Diag}(\Sigma|\nu \in N)} \in \mathcal{P}(\mathbb{R}^{N \times I}) | (\xi, \Sigma) \in L^I \times \mathcal{P}(I)). \]
The $N \times N$ orthogonal projection matrix given by $L = L_1 \times L_2$ is the given by

$$P \equiv \text{Diag}(P_1, P_2)$$

The condition for the existence of the ML estimator is $N \geq I + l$, where $l := \text{Dim}(L) = l_1 + l_2$. Since $N = N_1 + N_2$ it follows that the existence of the ML estimator under $H$ implies the existence of the ML estimator under $H_0$. The ML estimator at $x \in \mathbb{R}^{N \times I}$ under $H_0$ is thus

$$\hat{\xi}(x) = Px$$

and

$$\hat{\Sigma}(x) = \frac{1}{N}(x^t x - x^t Px)$$

with maximum value at $x \in \mathbb{R}^{N \times I}$

$$\frac{1}{\sqrt{2\pi}^{NI/2} |\hat{\Sigma}(x)|^{N/2}} \exp\{-\frac{NI}{2}\}.$$  

Note that

$$\hat{\xi}(x) = \text{Diag}(P_1, P_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} P_1 x_1 \\ P_2 x_2 \end{pmatrix} = \begin{pmatrix} \hat{\xi}_1(x_1) \\ \hat{\xi}_2(x_2) \end{pmatrix},$$

the estimator of $\xi \in L^I$ under $H_0$ expressed in terms of the estimators of $\xi_1 \in L_1^I$ and $\xi_2 \in L_2^I$ under $H$. Furthermore
expressing the ML estimator of $\Sigma \in P(I)$ under $H_0$ as in terms of the ML estimators of $\Sigma_1 \in P(I)$ and $\Sigma_2 \in P(I)$ under $H$. In fact $\hat{\Sigma}(x)$ is a weighted average of $\hat{\Sigma}_1(x_1)$ and $\hat{\Sigma}_2(x_2)$ with $N_1$ and $N_2$ as the weights.
The LR test statistics then becomes

\[
q(x) = \frac{1}{\sqrt{2\pi}^{NI/2} |\hat{\Sigma}(x)|^{N/2}} \exp\left\{-\frac{NI}{2}\right\} \frac{1}{\sqrt{2\pi}^{NI/2} |\hat{\Sigma}_1(x)|^{N_1/2} |\hat{\Sigma}_2(x)|^{N_2/2}} \exp\left\{-\frac{NI}{2}\right\} = \frac{|\hat{\Sigma}_1(x)|^{N_1/2} |\hat{\Sigma}_2(x)|^{N_2/2}}{|\hat{\Sigma}(x)|^{N/2}}.
\]

[ We have to mention that usually the so-called *unbiased* test statistic:

\[
u(x) = \frac{|\hat{\Sigma}_1(x)|^{(N_1-l_1)/2} |\hat{\Sigma}_2(x)|^{(N_2-l_2)/2}}{|\hat{\Sigma}(x)|^{(N-l)/2}}
\]
is used. But we stay with our principles and use the LR test statistics.]

The distribution of the equivalent test statistics

\[-2 \log(q(x)) = N \log(|\hat{\Sigma}(x)|) - N_1 \log(|\hat{\Sigma}_1(x)|) - N_2 \log(|\hat{\Sigma}_2(x)|)\]

with large values extreme can be approximated with a Chi-square distribution with \(f\) degrees of freedom, where \(f := [l_1 I + l_2 I + 2\frac{I(I+1)}{2}] - [lI + \frac{I(I+1)}{2}] = \frac{I(I+1)}{2}\).