APPENDIX A
DETAILS FROM CHAPTER 4

A.1. Definition of a Certain Orthogonal Matrix

\( \mathbb{R}^n \) can be partitioned into orthogonal subspaces as follows. Let \( \mathcal{K}_0 \) be such that \( \mathbb{R}^n = \mathcal{K}_0 \oplus \mathcal{L}(\widetilde{U}_1: \ldots: \widetilde{U}_p) \) and \( \mathcal{K}_0 \) is orthogonal to \( \mathcal{L}(\widetilde{U}_1: \ldots: \widetilde{U}_p) \). For \( s = 1, 2, \ldots, c-1 \) let \( \mathcal{K}_s \) be such that

\[ \mathcal{L}(\widetilde{U}_1: \ldots: \widetilde{U}_p) = \mathcal{K}_s \oplus \mathcal{L}(\widetilde{U}_{s+1}: \ldots: \widetilde{U}_p) \]

\( \mathcal{K}_s \) is orthogonal to \( \mathcal{L}(\widetilde{U}_{s+1}: \ldots: \widetilde{U}_p) \). Let \( \mathcal{K}_c = \mathcal{L}(\widetilde{U}_1: \ldots: \widetilde{U}_p) \). (The \( \widetilde{U}_s \) matrices are as defined in Section 1.3 and the partition of \( \{1, 2, \ldots, p\} \) into sets \( S_s \) is as described in Section 4.2.) Then there are \( c+1 \) mutually orthogonal vector spaces such that \( \mathbb{R}^n = \mathcal{K}_0 \oplus \mathcal{K}_1 \oplus \ldots \oplus \mathcal{K}_c \). Let the dimension of \( \mathcal{K}_s \) be \( \tilde{m}_s \) and let \( \mathcal{H}_{\tilde{s}} \) be an orthonormal basis for \( \mathcal{K}_s \). Then

\[ H_{\tilde{s}_1 \tilde{s}_2} \approx_{\tilde{s}_1 = \tilde{s}_2} \begin{cases} 0 & \text{if } \tilde{s}_1 \neq \tilde{s}_2 \\ \mathbb{I}_{\tilde{m}_{\tilde{s}_1}} & \text{if } \tilde{s}_1 = \tilde{s}_2 \end{cases} \]

Thus \( P = [H_{1: \tilde{H}_2: \ldots: \tilde{H}_c}] \) is an \( nn \) orthogonal matrix; that is, \( P'P = \mathbb{I} = \mathbb{I}P' \). Furthermore, for any \( s = 0, 1, \ldots, c \) and any \( i \in S_{s-1} : U_i^T H_s = \sim \) because the columns of \( H_{\tilde{s}} \) are orthogonal to all vectors in \( \mathcal{L}(\widetilde{U}_{s+1}: \ldots: \widetilde{U}_p) \). This yields

151
\begin{equation*}
\begin{aligned}
P_1 \begin{pmatrix} 
\sum_{i=0}^{p_1} b_i G_i \end{pmatrix}_{H_1} = 
\sum_{i=0}^{p_1} b_i U_i U_i^H H_1 
\end{aligned}
\end{equation*}

(\text{Let } U_0 \equiv \mathbf{I}_n.)

\begin{aligned}
&= \sum_{i=0}^{i_{s+1}-1} b_i U_i U_i^H H_1 \\
&= \sum_{i=0}^{s+1} b_i G_i H_i \\
&= \sum_{i=s+1}^{s+1} b_i G_i H_i .
\end{aligned}

Now since \( T = \sum_{i=0}^{n-1} \sigma_i G_i \) is positive definite, there exists a lower triangular matrix \( \hat{A} \) such that \( \Sigma_0 \equiv \hat{A} \hat{A}' \). But \( \hat{T}^t \Sigma \hat{T} \) is also positive definite and hence there exists \( \hat{T} \) upper triangular such that \( \hat{T}^t \Sigma \hat{T} = \hat{T}^t T \).

Then if \( \hat{T} = T^{-1} \), \( \hat{T} \) is also upper triangular and \( \Sigma \equiv A' PT \) is an \( n \times n \) matrix with the following properties.

1) \( \Sigma^t \Sigma = \hat{T}^t P^t A A' PT \),

\begin{align*}
&= \hat{T}^t P^t \Sigma_0 \Sigma P T \\
&= (\Sigma')^{-1} \hat{T}^t \Sigma \hat{T}^{-1} \\
&= \hat{T}.
\end{align*}

that is, \( \Sigma \) is orthogonal.
\[ Q = A'[H_0; H_1; \ldots; H_c] \left[ T_{00}; T_{01}; \ldots; T_{0c} \right. \]
\[ \left. Q : T_{11}; \ldots; T_{1c} \right. \]
\[ \vdots \]
\[ \left. 0 : 0 \right. \quad T_{cc} \]

\[ = A'[H^*_0; H^*_1; \ldots; H^*_c] \]
\[ = [Q_0; Q_1; \ldots; Q_c] \]

where \( Q_0 = A'[H^*_0] = A' \sum_{t=0}^{s} H_t T_{ts} \) and everything is partitioned so that all multiplications can be properly carried out.

Then it is true that

\[
\left( \sum_{i=0}^{P_1} b_{i \sim d_i} H^*_i \right)_{s+1}^{-1} = \left( \sum_{i=0}^{P_1} b_{i \sim d_i} U_i^' (\sum_{t=0}^{s} H_t T_{ts})_{i=s+1}^{-1} \right)

\[
= \left( \sum_{i=0}^{P_1} b_{i \sim d_i} U_i^' (\sum_{t=0}^{s} H_t T_{ts})_{i=s+1}^{-1} \right)

\[
= \left( \sum_{i=0}^{P_1} b_{i \sim d_i} H^*_i \right)_{s+1}^{-1}

\]

because \( H^*_i \) only involves \( H_t \) for \( t \leq s \). (For \( t \leq s \), \( i \in S^*_t \) implies \( i \in S^*_{t+1} \); thus \( G_t H^* = Q \). Therefore \( G_t H^* = Q \) for \( i \in S^*_{t+1} \).)
3) \( \mathcal{Q} \) is orthogonal so \( \mathcal{Q}' \mathcal{Q} = I \). This implies

\[
I = \begin{bmatrix}
\mathcal{Q}' \\
\mathcal{Q}_0' \\
\vdots \\
\mathcal{Q}_1' \\
\vdots \\
\mathcal{Q}_c'
\end{bmatrix} [\mathcal{Q}_0; \mathcal{Q}_1; \ldots; \mathcal{Q}_c]
\]

Therefore

\[
\mathcal{Q}_s \mathcal{Q}_t = \begin{cases}
\mathcal{Q}_s' & s \neq t \\
I_m & s = t
\end{cases}
\]


Let \( \mathcal{Q} = [\mathcal{Q}_0; \mathcal{Q}_1; \ldots; \mathcal{Q}_c] \) be defined as in Section A.1, where each \( \mathcal{Q}_s \) is \( n \times n_m \). Then if \( \mathcal{Q}_0 = \mathcal{A}' \) as usual, \( \mathcal{Z} = \mathcal{A}^{-1} (\mathcal{Y} - \mathcal{X}_0) \sim \mathcal{N}_n (\mathcal{Q}_0, \mathcal{I}) \) whenever \( \mathcal{Y} \sim \mathcal{N}_n (\mathcal{X}_0, \mathcal{I}) \). If \( \mathcal{w} \) is defined by \( \mathcal{w} = \mathcal{Q}' \mathcal{Z} \), then \( \mathcal{w} \sim \mathcal{N}_n (\mathcal{Q}_0, \mathcal{I}) \) also. But \( \mathcal{w} \) can be written \( \mathcal{w} = [\mathcal{w}_0', \mathcal{w}_1', \ldots, \mathcal{w}_c']' \), where
\[ w_s = Q_s' z = Q_s' A^{-1}(y - X \theta_s) \]. Each \( w_s \) is \( \tilde{m}_s x \) and \( w_s \sim \tilde{m}_s (Q, I_{\tilde{m}_s}) \) for \( s=0,1,\ldots,c \). Condition A.2.1 is defined as follows:

**CONDITION A.2.1.** For \( w_s \) as defined above, \( \frac{w_s'w_s}{\tilde{m}_s} \leq \frac{11}{10} \), \( s=0,1,\ldots,c \).

The following proposition concerning Condition A.2.1 is true.

**PROPOSITION A.2.1.** Under Assumptions 1.3.1-1.3.6 and 4.2.1-4.2.5, \( P[\text{Condition A.2.1 is true}] \rightarrow 1 \) as \( n \rightarrow \infty \).

**PROOF.**

\[ P[\text{Condition A.2.1 is not true}] \]

\[ = P\left\{ \frac{w_s' w_s}{\tilde{m}_s} > \frac{11}{10} \text{ for some } s=0,1,\ldots,c \right\} \]

\[ \leq \sum_{s=0}^{c} P\left\{ \frac{w_s' w_s}{\tilde{m}_s} > \frac{11}{10} \right\} \]

\[ = \sum_{s=0}^{c} P\left\{ \frac{w_s' w_s - \tilde{m}_s}{\tilde{m}_s} > \frac{1}{10} \right\} \]

\[ \leq \sum_{s=0}^{c} P\left\{ \left| \frac{w_s' w_s - \tilde{m}_s}{\tilde{m}_s} \right| > \frac{1}{10} \right\} . \]

But \( w_s' w_s \sim \chi^2_{\tilde{m}_s} \) and hence \( \delta(w_s' w_s) = \tilde{m}_s \) and \( \text{Var}(w_s' w_s) = 2\tilde{m}_s \) for \( s=0,1,\ldots,c \).
so by a simple application of Chebychev's Inequality

\[ P[\text{Condition A.2.1 is not true}] \leq \frac{c}{\sum_{s=0}^{200} \frac{\bar{m}_s}{\bar{n}_s}}. \text{ But each } \bar{m}_s \geq \nu_1 \text{ for some } i \in S, \text{ where } \nu_1 \text{ is as defined in Section 4.2. Thus each } \bar{m}_s \to \infty. \]

as \( n \to \infty \) by Assumptions 4.2.1 and 4.2.4. This implies that

\[ P[\text{Condition A.2.1 is not true}] \to 0 \text{ as } n \to \infty, \text{ which proves the proposition.} \]

A.3. Bounds for Various Inner Products and the Characteristic Roots of Various Matrices

In this section bounds for all the terms which must be dealt with in Section A.4 are found. These terms are either traces or inner products. The traces are of the form \( \text{tr} \ E \ G \ E \ G \ j \) and \( \text{tr} \ E \ G \ E \ G \ E \ G \ j \) for certain choices of \( E_1, E_2 \) and \( E_3 \); the inner products are of the form

\( E' \ X' \ E X E, E' \ X' \ E (X - \xi', Y' \xi), \text{ and } (X - \xi') \ E (X - \xi), \) where \( E \) equals

\( E \ G \ E_2, E \ G \ E_2 \ G \ E, \text{ or } E \ G \ E_2 \ G \ E \ G \ E, \) for various choices of \( E, E_2, E_3, E_4, E_5, E_6 \) and \( E_7. \) Bounds will be established for many characteristic roots leading up to bounding the desired inner products and traces.

Let \( b > 0 \) be given and let \( 0 < \delta < \frac{b}{2}. \) Let \( \gamma_1 \in S_b(\psi_0) \) and let \( \gamma_2 \in S_\delta(\gamma_1). \) Let \( \gamma_{1n} = (\beta', \tau_{a0}, \tau_{a1}, \ldots, \tau_{ap_1}'), \) \( a = 0, 1, 2 \) and recall that \( \gamma_1 \in S_b(\psi_0) \) implies \( ||\gamma_1 - \gamma_0|| < b \) which in turn implies that
\[ |\tau_{1i} - \tau_{0i}| < b \text{ for all } i=0,1,\ldots,p_1 \text{ and } |\beta_{1j} - \beta_{0j}| < b \text{ for all } j=1,2,\ldots,p_0. \] Similarly \[ |\tau_{1i} - \tau_{2i}| < \delta \text{ and } |\beta_{1j} - \beta_{2j}| < \delta. \] Recall that \[ \sigma_{0i} = \frac{\tau_{0i}}{n_i}. \] Let the following condition hold.

**CONDITION A.3.1.** For a fixed \( b > 0 \) and with \( \sigma_{0i} \) and \( n_i \) defined as in Chapters 1 and 4, \[ \frac{\sigma_{0i}}{2} > \frac{b}{n_i}, \quad i=0,1,\ldots,p_1. \]

It is always possible that Condition A.3.1 holds because the \( n_i \) are sequences increasing to infinity and all \( \sigma_{0i} \) are positive. Several consequences follow immediately from the above definitions.

**PROPOSITION A.3.1.** For \( x_{0n}, x_{1n}, x_{2n} \) as defined above, if Condition A.3.1 is true, then the following statements are true.

\[ (\beta_{0i} - \beta_{1i})' (\beta_{0i} - \beta_{1i}) \leq p_0 b^2, \]

\[ (\beta_{1i} - \beta_{2i})' (\beta_{1i} - \beta_{2i}) \leq p_0 \delta^2, \]

\[ 0 < \frac{\sigma_{01}}{2} < \frac{\tau_{1i}}{n_i} < \frac{3\sigma_{01}}{2}, \]

and

\[ 0 < \frac{\sigma_{01}}{4} < \frac{\tau_{2i}}{n_i} < \frac{7\sigma_{01}}{4}, \quad i=0,1,\ldots,p_1. \]
PROOF.

\[(\beta_0 - \beta_1)'(\beta_0 - \beta_1) = \sum_{j=1}^{p_0} (\beta_{0j} - \beta_{1j})^2 \leq p_0 \sigma^2.\]

\[(\beta_{1} - \beta_{2})'(\beta_{1} - \beta_{2}) = \sum_{j=1}^{p_0} (\beta_{1j} - \beta_{2j})^2 \leq p_0 \sigma^2.\]

\[
\frac{\tau_{11}}{n_1} = \frac{\tau_{0i}}{n_1} + \frac{\tau_{11} - \tau_{0i}}{n_1},
\]

\[
= \sigma_{0i} + \frac{\tau_{11} - \tau_{0i}}{n_1},
\]

\[
\leq \sigma_{0i} + \frac{b}{n_1} < \frac{3\sigma_{0i}}{2},
\]

and

\[
\frac{\tau_{11}}{n_1} \geq \sigma_{0i} - \frac{b}{n_1} > \frac{\sigma_{0i}}{2} > 0.
\]

Similarly

\[
\frac{\tau_{21}}{n_1} = \frac{\tau_{0i}}{n_1} + \frac{\tau_{11} - \tau_{0i}}{n_1} + \frac{\tau_{21} - \tau_{11}}{n_1},
\]

\[
\leq \sigma_{0i} + \frac{b}{n_1} + \frac{\delta}{n_1} < \sigma_{0i} + \frac{\sigma_{0i}}{2} + \frac{\sigma_{0i}}{4} = \frac{7\sigma_{0i}}{4},
\]

and

\[
\frac{\tau_{21}}{n_1} \geq \sigma_{0i} - \frac{b}{n_1} - \frac{\delta}{n_1} > \frac{\sigma_{0i}}{4} > 0.
\]
Now define $T_a = \sum_{i=0}^{p_1} \frac{\tau_{ai}}{n_i} G_i$, $a=0,1,2$ and recall $T_0 = \Sigma_0$.

Proposition A.3.1 is used to prove the following proposition.

**PROPOSITION A.3.2.** If Condition A.3.1 is true, the following statements are true.

\[
\lambda_{\text{max}}(\Sigma_0^{-1} T_0) \leq \frac{1}{\sigma_{01}}, \quad \lambda_{\text{max}}(\Sigma_0^{-1} T_1) \leq \frac{3}{2}, \quad \lambda_{\text{max}}(\Sigma_0^{-1} T_2) \leq \frac{7}{4}, \quad \lambda_{\text{max}}(T_1^{-1} T_0) \leq 2,
\]

\[
\lambda_{\text{max}}(T_2^{-1} T_0) \leq 4, \quad \lambda_{\text{max}}(T_2^{-1} T_1) \leq \frac{7}{2}, \quad \lambda_{\text{max}}(T_2^{-1} T_1) \leq 6,
\]

\[
\max_{k=1,2,\ldots,n} |\lambda_k[\Sigma_0^{-1}(T_1 T_2)]| = \max_{k=1,2,\ldots,n} |\lambda_k[\Sigma_0^{-1}(T_2 T_1)]| \leq \frac{\delta}{\min_{i=0,1,\ldots,p_1} (n_i \sigma_{0i})},
\]

\[
\max_{k=1,2,\ldots,n} |\lambda_k[\Sigma_0^{-1}(T_0 T_2)]| = \max_{k=1,2,\ldots,n} |\lambda_k[\Sigma_0^{-1}(T_2 T_0)]| \leq \frac{b}{\min_{i=0,1,\ldots,p_1} (n_i \sigma_{0i})},
\]

\[
\max_{k=1,2,\ldots,n} |\lambda_k[\Sigma_0^{-1}(T_0 T_1)]| = \max_{k=1,2,\ldots,n} |\lambda_k[\Sigma_0^{-1}(T_1 T_2)]| \leq \frac{2b}{\min_{i=0,1,\ldots,p_1} (n_i \sigma_{0i})},
\]

\[
\max_{k=1,2,\ldots,n} |\lambda_k[\Sigma_0^{-1}(T_1 T_1)]| = \max_{k=1,2,\ldots,n} |\lambda_k[\Sigma_0^{-1}(T_2 T_2)]| \leq \frac{2\delta}{\min_{i=0,1,\ldots,p_1} (n_i \sigma_{0i})},
\]

\[
\max_{k=1,2,\ldots,n} |\lambda_k[\Sigma_0^{-1}(T_2 T_1)]| = \max_{k=1,2,\ldots,n} |\lambda_k[\Sigma_0^{-1}(T_0 T_2)]| \leq \frac{4\delta}{\min_{i=0,1,\ldots,p_1} (n_i \sigma_{0i})}.
\]
PROOF.

Continually apply Lemma B.7 to obtain

\[ \lambda_{\text{max}} (\tau^{-1}_{0} G_{i}) \leq \max(0, \frac{1}{\sigma_{0i}}) = \frac{1}{\sigma_{0i}} \]

since \( G_{i} = \sum_{j=0}^{P_{1}} o_{i} g_{j} + l_{i} g_{i} \).

\[ \lambda_{\text{max}} (\tau^{-1}_{0} T_{1}) \leq \max_{i=0,1,\ldots,P_{1}} \frac{\tau_{1i}/n_{i}}{\sigma_{0i}} \leq \frac{3}{2} . \]

\[ \lambda_{\text{max}} (\tau^{-1}_{0} T_{2}) \leq \max_{i=0,1,\ldots,P_{1}} \frac{\tau_{2i}/n_{i}}{\sigma_{0i}} \leq \frac{7}{4} . \]

\[ \lambda_{\text{max}} (\tau^{-1}_{1} T_{0}) \leq \max_{i=0,1,\ldots,P_{1}} \frac{\sigma_{0i}}{\tau_{1i}/n_{i}} \leq 2 . \]

\[ \lambda_{\text{max}} (\tau^{-1}_{2} T_{0}) \leq \max_{i=0,1,\ldots,P_{1}} \frac{\sigma_{0i}}{\tau_{2i}/n_{i}} \leq 4 . \]

\[ \lambda_{\text{max}} (\tau^{-1}_{2} T_{1}) \leq \max_{i=0,1,\ldots,P_{1}} \frac{\tau_{2i}/n_{i}}{\tau_{1i}/n_{i}} \leq \frac{7}{2} . \]

\[ \lambda_{\text{max}} (\tau^{-1}_{0} T_{2}) \leq \max_{i=0,1,\ldots,P_{1}} \frac{\tau_{1i}/n_{i}}{\tau_{2i}/n_{i}} \leq 6 . \]
For the remaining inequalities

\[
\max_{k=1,2,\ldots,n} |\lambda_k [\Sigma_0^{-1}(\tau_1 - \tau_2)]| = \max_{k=1,2,\ldots,n} \left| \frac{x_k' (\tau_1 - \tau_2) x_k}{x_k' \Sigma x_k} \right|
\]

by Lemma B.6,

\[
\max_{k=1,2,\ldots,n} \left| \frac{p_1 \Sigma \left( \frac{\tau_{li} - \tau_{2i}}{\Sigma} \right) x_k' g_i x_k}{n_i} \right| = \max_{k=1,2,\ldots,n} \left| \frac{p_1 \Sigma \sigma_{0i} x_k' g_i x_k}{n_i} \right|
\]

because \( \sigma_{0i} > 0 \) and \( x_k' g_i x_k \geq 0 \),

\[
\max_{k=1,2,\ldots,n} \left| \lambda_k [\Sigma_0^{-1}(\tau_1 - \tau_2)]| \right| \leq \max_{k=1,2,\ldots,n} \left| \frac{p_1 \Sigma \left( \frac{\tau_{li} - \tau_{2i}}{\Sigma} \right) x_k' g_i x_k}{n_i} \right|
\]

\[
\leq \max_{k=1,2,\ldots,n} \left| \frac{p_1 \Sigma \sigma_{0i} x_k' g_i x_k}{n_i} \right|
\]

\[
\leq \max_{i=0,1,\ldots,p} \left| \frac{p_1 \Sigma \sigma_{0i} x_k' g_i x_k}{n_i} \right|
\]
\[
\max_{k=1,2,\ldots,n} |\lambda_k[\Sigma_0^{-1}(T_1 - \Sigma_0)]| \leq \frac{\delta}{\min_{i=0,1,\ldots,p_1} (n_i \sigma_{0i})}
\]

by Lemma B.3. Similarly,

\[
\max_{k=1,2,\ldots,n} |\lambda_k[\Sigma_0^{-1}(T_1 - \Sigma_0)]| = \max_{i=0,1,\ldots,p_1} \frac{|\tau_{1i}/n_i - \sigma_{0i}|}{\sigma_{0i}}
\]

\[
= \max_{i=0,1,\ldots,p_1} \frac{|\tau_{1i} - \sigma_{0i}|}{n_i \sigma_{0i}}
\]

\[
\leq \frac{b}{\min_{i=0,1,\ldots,p_1} (n_i \sigma_{0i})}.
\]

\[
\max_{k=1,2,\ldots,n} |\lambda_k[T_1^{-1}(T_1 - \Sigma_0)]| \leq \max_{i=0,1,\ldots,p_1} \frac{|\tau_{1i}/n_i - \sigma_{0i}|}{\tau_{1i}/n_i}
\]

\[
\leq \frac{2b}{\min_{i=0,1,\ldots,p_1} (n_i \sigma_{0i})}.
\]

\[
\max_{k=1,2,\ldots,n} |\lambda_k[T_1^{-1}(T_1 - \Sigma_2)]| \leq \max_{i=0,1,\ldots,p_1} \frac{|(\tau_{1i} - \tau_{2i})/n_i|}{\tau_{1i}/n_i}
\]
\[
\leq \max_{i=0,1,\ldots,p_1} \frac{\delta/n_i}{\sigma_{0i}/2}
\]

\[
\leq \frac{2\delta}{\min_{i=0,1,\ldots,p_1} (n_i \sigma_{0i})}.
\]

\[
\max_{k=1,2,\ldots,n} |\lambda_k(T_2^{-1}(T_1 - T_2))| \leq \max_{i=0,1,\ldots,p_1} \frac{|(\tau_{1i} - \tau_{2i})/n_i|}{\tau_{2i}/n_i}
\]

\[
\leq \frac{4\delta}{\min_{i=0,1,\ldots,p_1} (n_i \sigma_{0i})}.
\]

**Proposition A.3.3.** Let \(E_1, E_2, E_3\) and \(E_4\) be any \(n \times n\) symmetric matrices and \(\Sigma_0 = AA'\); then if Condition A.3.1 is true the following statements are true.

\[
\lambda_{\max}(A' E_1 E_2 E_3 A) \leq \max_{k=1,2,\ldots,n} \left| \lambda_k(\Sigma_0 E_1) \right|^2 \max_{k=1,2,\ldots,n} \left| \lambda_k(\Sigma_0 E_2) \right|^2 \frac{1}{\sigma_{0i}^2},
\]

\[
\lambda_{\max}(A' E_1 E_2 E_3 A' E_4 E_2 E_3 A).
\]
\[ \lambda_{\text{max}} \left( A' E G E A A' E G E A \right) \leq \max_{\ell=1,2,\ldots,n} \left| \lambda_{\ell}(\Sigma_{0\ell}) \right|^2 \max_{\ell=1,2,\ldots,n} \left| \lambda_{\ell}(\Sigma_{0\ell-2}) \right|^2 \max_{\ell=1,2,\ldots,n} \left| \lambda_{\ell}(\Sigma_{0\ell-3}) \right|^2 \left( \frac{1}{\sigma_{01}\sigma_{0j}\sigma_{0k}} \right) \]

**Proof.**

The proof is given for the first case only; the other cases are proved analogously.

\[ \lambda_{\text{max}} \left( A' E G E A A' E G E A \right) \]

\[ = \lambda_{\text{max}} \left( A' E A A' E G E A A A' E G E A^{-1} - A' A A' E G A^{-1} \right) \]

\[ \leq \lambda_{\text{max}} \left( A' E A A' E G A^{-1} \right) \lambda_{\text{max}} \left( A' A A' E G A^{-1} \right) \lambda_{\text{max}} \left( A' E G E A A' E G A^{-1} \right) \]
by two applications of Lemma B.8,

\[ \lambda_{\text{max}}(\Sigma_{0}^{-1} \Pi_{1})^2 \lambda_{\text{max}}(\Sigma_{0}^{-1} \Pi_{2})^2 \lambda_{\text{max}}(\Sigma_{0}^{-1} \Pi_{2})^2 \]

by Lemma B.9,

\[ \leq \max_{k=1,2,\ldots,n} |\lambda_{k}(\Sigma_{0}^{-1} \Pi_{1})|^2 \max_{k=1,2,\ldots,n} |\lambda_{k}(\Sigma_{0}^{-1} \Pi_{2})|^2 \frac{1}{\sigma_{0}^{-1}} \]

by Lemmas B.14 and B.7. |||

Proposition A.3.4 deals with inner products.

PROPOSITION A.3.4. Let \( \xi_1 \) and \( \xi_2 \) be any \( p \times 1 \) vectors and \( F_0, F_1, \) and \( F_2 \)
any \( n \times n \) matrices; then the following statements are true.

\[ (\xi'_{1} F_{1} \xi_{2})^2 \leq (\xi'_{1} F_{1} F_{2}^{-1} F_{1} \xi_{2})^2 \lambda_{\text{max}}^{2}(X^{'T} X_{0}^{-1} X)^{2} \lambda_{\text{max}}(A^{'} F_{1}^{-1} A^{-1} F_{1} A), \]

\[ (\xi'_{1} X_{0}^{-1} X_{0}^{-1} X_{0})^2 \leq (\xi'_{1} X_{0}^{-1} X_{0})^2 \lambda_{\text{max}}(X^{'T} X_{0}^{-1} X)^{2} \lambda_{\text{max}}(A^{'} F_{1}^{-1} A^{-1} F_{1} A), \]

\[ (X-X_{0}^{'} F_{1}^{'} A^{-1} F_{2} (X-X_{0})) \]

\[ (X-X_{0})^{'} F_{2} F_{1} F_{2}^{-1} F_{1} (X-X_{0}) \leq \lambda_{\text{max}}(A^{'} F_{1}^{-1} A^{-1} F_{1} A)(X-X_{0})^{'} F_{2} F_{1} F_{2}^{-1} F_{1} (X-X_{0}), \]

\[ (X-X_{0})^{'} F_{2} F_{1} F_{2}^{-1} F_{1} (X-X_{0}) \leq \lambda_{\text{max}}(A^{'} F_{1}^{-1} A^{-1} F_{1} A)(X-X_{0})^{'} F_{2} F_{1} F_{2}^{-1} F_{1} (X-X_{0}), \]
PROOF.

\[(\varepsilon_{1}^2 \varepsilon_{2}^2 \bar{A}^2 \bar{A}^2)^2 = (\varepsilon_{1}^2 \varepsilon_{2}^2 \bar{A}^2 \bar{A}^2)^2 \]

\[\leq (\varepsilon_{1}^2 \varepsilon_{2}^2)(\varepsilon_{1}^2 \varepsilon_{2}^2)\lambda_{\max}(\bar{A}^2 \bar{A}^2)\]

by Lemma B.12,

\[\leq (\varepsilon_{1}^2 \varepsilon_{2}^2)(\varepsilon_{1}^2 \varepsilon_{2}^2)\lambda_{\max}(\bar{A}^2 \bar{A}^2)\]

by definition of characteristic root. The other cases are proved analogously. |||

Now \(\bar{A}^2 \bar{A}^2\) of Proposition A.3.1 will be of the correct form to plug into Proposition A.3.3. It remains to bound terms of the form \((\bar{X} - \bar{X}_0)^2\bar{A}^2 \bar{A}^2(\bar{X} - \bar{X}_0)\). This is done using Condition A.2.1. Let \(Q\) be defined as in Section A.1 and let \(\bar{y}\) be as defined in Section A.2. Then

\[\bar{X} - \bar{X}_0 = \bar{A}^{-1}(\bar{X} - \bar{X}_0)\]

\[= \bar{A}Q\bar{A}^{-1}(\bar{X} - \bar{X}_0)\]

\[= \bar{A}Q\bar{y}\]
\[
\begin{pmatrix}
    x_0 \\
    x_1 \\
    \vdots \\
    x_c \\
\end{pmatrix} = A \sum_{s=0}^{c} Q_s W_s.
\]

This yields
\[
(\mathbf{x} - \bar{x}_0) F_{\mathcal{E}_c} A^{-t} F_{\mathcal{E}_c} (\mathbf{x} - \bar{x}_0)
\]

\[
= \sum_{t=0}^{c} \sum_{s=0}^{c} W'_s Q'A'F'A^{-t} A^{-1} F_{\mathcal{E}_c} A Q_s W_s.
\]

The Cauchy-Schwarz inequality gives a bound for each term as
\[
(W'_s Q'A'F'A^{-t} A^{-1} F_{\mathcal{E}_c} A Q_s W_s)^2
\]

\[
\leq (W'_s Q'A'F'A^{-t} A^{-1} F_{\mathcal{E}_c} A Q_s W_s) (W'_s Q'A'F'A^{-t} A^{-1} F_{\mathcal{E}_c} A Q_s W_s).
\]

But
\[
W'_s Q'A'F'A^{-t} A^{-1} F_{\mathcal{E}_c} A Q_s W_s \leq \lambda_{\max} (Q'_s A'T_A^{-t} A^{-1} F_{\mathcal{E}_c} A Q_s)
\]

and
\[
W'_s W_s \leq \frac{11}{10} \tilde{m}_s \text{ by Condition A.2.1.} \text{ Thus bounds are required on}
\]
\[ \lambda_{\text{max}}(Q_s^A F_s^{A^t} A^{-1} F_s A Q_s^A) \] for various choices of $F_s$. These are provided by the next three propositions.

**PROPOSITION A.3.5.** If \( F_s = G_s F_s^{-1} \) and if Condition A.3.1 is true then the following statements are true.

- If \( i \notin S_{s+1}^* \), \[ \lambda_{\text{max}}(Q_s^A F_s^{A^t} A^{-1} F_s A Q_s^A) \leq \frac{1}{2c_{01}}. \]

- If \( i \in S_{s+1}^* \), \[ \lambda_{\text{max}}(Q_s^A F_s^{A^t} A^{-1} F_s A Q_s^A) = 0. \]

**PROOF.**

Recall \( S_{s+1}^* = \{ i_{s+1}, i_{s+1} + 1, \ldots, p_1 \} \). Then

\[
\lambda_{\text{max}}(Q_s^A F_s^{A^t} A^{-1} F_s A Q_s^A) = \lambda_{\text{max}}(Q_s^A F_s^{-1} G_s A^{-1} G_s A^{-1} A^{-1} G_s A^{-1} A^{-1} Q_s^A)
\]

\[
= \lambda_{\text{max}}(Q_s^A F_s^{-1} G_s A^{-1} G_s A^{-1} Q_s^A)
\]

\[
= \sup_{\gamma \neq 0} \frac{\gamma ' Q_s^A F_s^{-1} G_s A^{-1} G_s A^{-1} Q_s^A \gamma}{\gamma ' G_s A^{-1} G_s A^{-1} Q_s^A \gamma}
\]

\[
= \sup_{\gamma \neq 0} \frac{\gamma ' Q_s^A \gamma}{\gamma ' G_s A^{-1} G_s A^{-1} Q_s^A \gamma}
\]

\[
= \sup_{\gamma \neq 0} \frac{\gamma ' Q_s^A \gamma}{\gamma ' G_s A^{-1} G_s A^{-1} Q_s^A \gamma}
\]
because \( Q_s'Q_s = I \),

\[
\sup_{x \neq 0} \frac{x' A^{-1} g_s A^{-t} \frac{1}{\frac{1}{\lambda}} \frac{1}{\lambda} A^{-t} x'}{x' x'} = \frac{1}{\lambda_{\max}(s_0, A)} \leq \frac{1}{\sigma_{01}}.
\]

However, if \( i \in S_{s+1} \) then \( g_s A^{-t} q_s = g_h A^{-t} = 0 \) as was shown in Section A.1.

Then the matrix in question is the zero matrix and hence has characteristic roots all zero. 

PROPOSITION A.3.6. If \( \tilde{P} = (\Sigma_0 - T_1)^{-1} \Sigma_0 \) where \( T_1 \) is as above and if Condition A.3.1 is true then the following statement is true.

\[
\lambda_{\max}(Q_s' A' T_1^{-1} A^{-t} F_s A Q_s) \leq \frac{b^2}{\min_{i=0,1,\ldots,i_{s+1}-1} (n_i \sigma_{0i})^2}.
\]

PROOF.

\[
\lambda_{\max}(Q_s' A' T_1^{-1} A^{-t} F_s A Q_s) = \lambda_{\max}(Q_s' A^{-1}(\Sigma_0 - T_1) A^{-t} (\Sigma_0 - T_1) A^{-t} Q_s).
\]

But

\[
(T_1 - \Sigma_0) A^{-t} Q_s = \left( \sum_{i=0}^{p_1} \frac{r_{0i} - r_{11}}{n_i} g_i \right) H_s^*.
\]
\[
\begin{align*}
&= \left( \sum_{i=0}^{i_{s+1}-1} \frac{(\tau_{0i} - \tau_{1i})}{n_i} g_0 \right) h_s^* \\
&= F^* h_s^*
\end{align*}
\]

\[
= F^* A^{-t} h_s^*
\]

\[
= F^* A^{-t} Q_s
\]

where

\[
F^* = \sum_{i=0}^{i_{s+1}-1} \frac{(\tau_{0i} - \tau_{1i})}{n_i} g_0
\]

which is symmetric. Therefore

\[
\lambda_{\text{max}}(Q_{A-1}^* F^* A^{-t} A^{-1} F A Q_{\sim}) = \lambda_{\text{max}}(Q_{A-1}^* F^* A^{-t} A^{-1} F A^{-t} Q_{\sim})
\]

\[
= \sup_{\gamma \in \mathcal{O}} \frac{\chi' Q_{A-1}^* F^* A^{-t} A^{-1} F A^{-t} \chi}{\chi' \chi}
\]

\[
\leq \sup_{x \in \mathcal{O}} \frac{\chi' A^{-1} F A^{-t} A^{-1} F A^{-t} \chi}{\chi' \chi}
\]

as in Proposition A.3.5,

\[
= \max_{k=1,2,\ldots,n} |\lambda_k(\Sigma_0^* F^*)|^2
\]
\[ \lambda_{\text{max}}(Q' A' F'A^{-1} F^T A Q) \leq \frac{\delta^2}{\min_{i=0,1,\ldots,i_{s+1}-1} \left( n_i \sigma_{0i} \right)^2} \]

by Lemma B.7 just as in Proposition A.3.2. \|

PROPOSITION A.3.7. If \( \overline{F}_2 \) = \((T_1 - T_2) \Sigma_0^{-1}\) where \( T_1 \) and \( T_2 \) are as above and if Condition A.3.1 is true then the following statement is true.

\[ \lambda_{\text{max}}(Q' A' F'A^{-1} F^T A Q) \leq \frac{\delta^2}{\min_{i=0,1,\ldots,i_{s+1}-1} \left( n_i \sigma_{0i} \right)^2} \]

PROOF.

Proceed exactly as in the proof of Proposition A.3.6 but at last step, instead of \(|\tau_{01} - \tau_{11}| < b\) use \(|\tau_{11} - \tau_{21}| < \delta\). \|

There are two different types of \( \overline{F}_2 \) which will be encountered. However, the necessary bounds reduce to those of the form above as will be seen. The first different \( \overline{F}_2 \) is \( \overline{F}_2 = G \Sigma^{-1}_{A2}. \) But

\[ \Sigma_{A1}^{-1} = \Sigma_0^{-1} + \Sigma^{-1}_{A2} - \Sigma_0^{-1} \]

\[ = \Sigma_0^{-1} + \Lambda \]

where \( \Lambda = \Sigma_{A1}^{-1} - \Sigma_0^{-1} = \Sigma_0^{-1}(\Sigma_0 - \Sigma_{A2}) \Sigma^{-1}_0 = \Sigma_{A1}^{-1}(\Sigma_0 - \Sigma_{A2}) \Sigma_0^{-1}. \) Therefore

\[ T_{11}^{-1} B T_{11}^{-1} = \Sigma_0^{-1} B \Sigma_0^{-1} + \Lambda B \Lambda + \Sigma_0^{-1} B \Lambda + \Lambda B \Sigma_0^{-1} \] for any \( B. \) In this case
\[ B = G_1 A^{-1} A^{-1} G_1 \] and so

\[ \chi' \left( \bar{\gamma}' A_t F^* A_t^{-1} \right) \bar{\gamma} = \chi' \left( \bar{\gamma}' A_t^{-1} G_1 A_t^{-1} G_1 G_1^{-1} A_0 \right) \bar{\gamma} \]

\[ = \chi' \left( \bar{\gamma} A^{-1} G_1 A^{-1} \right) \bar{\gamma}^{-1} G_1 G_1^{-1} A_0, \]

\[ + \chi' \left( \bar{\gamma}' (T_0) T^{-1} G_1 A^{-1} G_1^{-1} T_0^{-1} (T_0) \right) \bar{\gamma}^{-1} A_0, \]

\[ + \chi' \left( \bar{\gamma}' (T_0) T^{-1} G_1 A^{-1} G_1^{-1} T_0^{-1} (T_0) \right) \bar{\gamma}^{-1} A_0, \]

\[ + \chi' \left( \bar{\gamma}' (T_0) T^{-1} G_1 A^{-1} G_1^{-1} T_0^{-1} (T_0) \right) \bar{\gamma}^{-1} A_0, \]

The first term is exactly as in Proposition A.3.5. The second term by the same reasoning as Propositions A.3.6 and A.3.4 is less than

\[ \chi' \left( \gamma_{\lambda_{\max}} (T_0^{-1} G_1) \right) \lambda_{\max}^2 (T_0^{-1} T_0) \max_{k=1,2,...,n} \left| \lambda_k (T_0^{-1} T_0^*) \right|^2. \]

The third and fourth terms are equal and by one application of the Cauchy-Schwarz Inequality, their squares are bounded by the product of the first two terms. Thus bounds exist for all terms based on Propositions A.3.5 and A.3.6.
The second different term is $E_2 = (T_1 - T_2)^{t-1}$. This yields a decomposition as above with $B = (T_1 - T_2)^{t-1}$. As above the squares of the third and fourth terms will be bounded by the product of the first two terms. The first term will be $\chi' \gamma' A^t Z_0^{-1}(T_1 - T_2)^{t-1} A^{-t} Z_0^{-1}(T_1 - T_2) E_0 Z_0^{-1} Z_0^{-1}$, which is exactly the same term that is bounded in Proposition A.3.7. The second term will be $\chi' \gamma' A^t Z_0^{-1}(T_1 - T_2)^{t-1} A^{-t} Z_0^{-1}(T_1 - T_2) E_0 Z_0^{-1} Z_0^{-1}$; it is easily seen that this term will be bounded by

$$\chi' \gamma' \frac{\max_k |\lambda_k(Z_0^{-1}(T_1 - T_2))|^2}{\max_k |\lambda_k(Z_0^{-1} Z_0^{-1})|^2}.$$ Thus all terms are bounded by previous propositions.

Bounds have now been found for all necessary inner products; bounds for traces are covered by the next two propositions.

PROPOSITION A.3.8. If $E_1$, $E_2$, and $E_3$ are any n×n matrices, then the following statements are true.

$$|\text{tr} E_1 G E_2 G| \leq \min(m_1, m_j) \max_k |\lambda_k(Z_0^{-1} Z_0^{-1})| \frac{1}{\sigma_0 \sigma_0}$$

$$\frac{|\text{tr} E_1 G E_2 G E_3 G|}{\sigma_0} \leq \min(m_1, m_2, m_k) \max_k |\lambda_k(Z_0^{-1} Z_0^{-1})| \frac{1}{\sigma_0 \sigma_0 \sigma_0}.$$}

PROOF.

The proof is given for the first case; the second is proved analogously. $E_1 G E_2 G$ has rank at most $\min(m_1, m_2)$ and hence has at
most \( \min(m_i, m_j) \) nonzero characteristic roots.

\[
\left| \text{tr} E_{\tilde{1}_{d_i} \tilde{2}_{d_j}} \right| = \sum_{k=1}^{n} \lambda_k(E_{\tilde{1}_{d_i} \tilde{2}_{d_j}}) \\
\leq \sum_{k=1}^{n} |\lambda_k(E_{\tilde{1}_{d_i} \tilde{2}_{d_j}})| \\
\leq \min(m_i, m_j) \max_{k=1,2,...,n} |\lambda_k(E_{\tilde{1}_{d_i} \tilde{2}_{d_j}})|.
\]

But

\[
\max_{k=1,2,...,n} |\lambda_k(E_{\tilde{1}_{d_i} \tilde{2}_{d_j}})| = \max_{k=1,2,...,n} |\lambda_k(U_j^{\prime} A_t^{-1} U_j)|
\]

by Lemma B.11 and \( G_j = U_j U_j^{\prime} \),

\[
\leq \lambda_{\max}(U_j^{\prime} A_t^{-1} U_j) \max_{k=1,2,...,n} |\lambda_k(A_{E_{\tilde{1}_{d_i}} E_{\tilde{2}_{d_j}}})|
\]

by Lemma B.8,

\[
\leq \lambda_{\max}(\Sigma_0^{-1} G_j) \max_{k=1,2,...,n} |\lambda_k(U_j^{\prime} A_{E_{\tilde{1}_{d_i}} E_{\tilde{2}_{d_j}}} A_{E_{\tilde{2}_{d_j}}}^{-1} U_j)|
\]

by Lemma B.11 again,

\[
\leq \lambda_{\max}(\Sigma_0^{-1} G_j) \lambda_{\max}(\Sigma_0^{-1} G_j) \max_{k=1,2,...,n} |\lambda_k(A_{E_{\tilde{2}_{d_j}} A_{E_{\tilde{1}_{d_i}}} A_{E_{\tilde{2}_{d_j}}}})|
\]

by Lemma B.8 again,

\[
\leq \max_{k=1,2,...,n} \left| \frac{\lambda_k(A_{E_{\tilde{2}_{d_j}} A_{E_{\tilde{1}_{d_i}}} A_{E_{\tilde{2}_{d_j}}}})}{\sigma_{0i} \sigma_{0j}} \right|
\]

by Proposition A.3.2. |||
The reason that $\lambda_k(A E_1 A A E_2 A)$ cannot be simplified further is that in general $\max|\lambda_k(AB)| \leq \max|\lambda_k(A)| \max|\lambda_k(B)|$. However, the choices of $E_1$, $E_2$ and $E_3$ needed in Chapter 4 are always of a form that allows simplification. In particular, Proposition A.3.9 is true.

**PROPOSITION A.3.9.** If $T_1, T_2, A$ are as above and if Condition A.3.1 is true then for the following choices of $E_1$ and $E_2$, the cited bounds result for $\max_{k=1,2,\ldots,n} |\lambda_k(A E_1 A A E_2 A)|$.

\[
\begin{align*}
E_1 & \quad E_2 \\
\Sigma_0 & \quad \Sigma_0 \\
T_1^{-1} & \quad T_1^{-1} \\
T_1^{-1}(T_1^{-1} T_2) T_2^{-1} & \quad T_1^{-1}(T_1^{-1} T_2) T_2^{-1} \\
T_1^{-1} T_2^{-1} & \quad T_1^{-1} T_2^{-1} \\
T_1^{-1}(T_1^{-1} T_2) T_2^{-1} & \quad T_1^{-1}(T_1^{-1} T_2) T_2^{-1} \\
\end{align*}
\]

**Bound for $\max|\lambda_k(A E_1 A A E_2 A)|$**

\[
\begin{align*}
1 & \quad \frac{2b}{\min_{i=0,1,\ldots,P_1} (n_i \sigma_{oi})} \\
\frac{4b^2}{\min_{i=0,1,\ldots,P_1} (n_i \sigma_{oi})^2} & \quad \frac{166}{\min_{i=0,1,\ldots,P_1} (n_i \sigma_{oi})} \\
\frac{64b^2}{\min_{i=0,1,\ldots,P_1} (n_i \sigma_{oi})^2} & \quad \frac{\min_{i=0,1,\ldots,P_1} (n_i \sigma_{oi})}{(n_i \sigma_{oi})^2}
\end{align*}
\]
The following choices of $E_1$, $E_2$ and $E_3$ yield the cited bounds for

$$\max_{k=1,2,...,n} |\lambda_k (A'E_1A + E_2A'EA) A|.$$

<table>
<thead>
<tr>
<th>$E_1$</th>
<th>$E_2$</th>
<th>$E_3$</th>
<th>Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{11}$</td>
<td>$T_{12}$</td>
<td>$T_{13}$</td>
<td>$8$</td>
</tr>
<tr>
<td>$T_{21}$</td>
<td>$T_{22}$</td>
<td>$T_{23}$</td>
<td>$\frac{32 \delta}{\min_{i=0,1,...,p_1} (n_i \sigma_{0i})}$</td>
</tr>
<tr>
<td>$T_{31}$</td>
<td>$T_{32}$</td>
<td>$T_{33}$</td>
<td>$\frac{128 \delta^2}{\min_{i=0,1,...,p_1} (n_i \sigma_{0i})^2}$</td>
</tr>
<tr>
<td>$T_{41}$</td>
<td>$T_{42}$</td>
<td>$T_{43}$</td>
<td>$\frac{512 \delta^3}{\min_{i=0,1,...,p_1} (n_i \sigma_{0i})^3}$</td>
</tr>
</tbody>
</table>

PROOF.

The first case is obvious. The second case simplifies to

$$\max |\lambda_k [T_{11}^{-1}(T_{20} - T_{21})]|$$

which is bounded by Proposition A.3.2. In cases three and four Lemma B.14 applies as it does in the sixth case. The fifth case is proved here for illustrative purposes.

$$\max_{k=1,2,...,n} |\lambda_k [(A'E_1A + E_2A'EA) A]| = \max_{k=1,2,...,n} |\lambda_k [A' (T_{12} - T_{22}) T_{22}^{-1} A]|$$

by Lemma B.11,

$$\leq \lambda_{\max} (A' T_{22}^{-1}) \max_{k=1,2,...,n} |\lambda_k [(T_{12} - T_{22}) T_{22}^{-1}]|$$
by Lemma B.13 and B.1k,

\[ \frac{h_j}{\min(n_i \sigma_{0i})} \]

by Proposition A.3.2. Similar logic applies to the second set of bounds. Both sets of bounds are unchanged upon permutations of \( E_1 \), \( E_2 \) and \( E_3 \). |||

This section concludes with some remarks which firm up the bounds obtained in the rest of the section.

PROPOSITION A.3.10. There exists some constant \( B \) such that the following statements are true.

\[ \frac{1}{n_{d_1+1}} \lambda_{\max}(X' Z_0 X) \leq B; \quad \frac{\mu_s}{\min n_i^2} \leq B \quad \text{and} \quad \frac{\mu_{E_s}}{\min(n_i \sigma_{0i})^2} \leq B, \]

\[ i=0, 1, \ldots, i_{s+1}-1 \]

\[ s=0, 1, \ldots, c; \quad \frac{\min(m_i, m_j)}{n_i n_j} \leq B \quad \text{and} \quad \frac{\min(m_i, m_j, m_k)}{n_i n_j n_k} \rightarrow 0 \quad i, j, k=0, 1, \ldots, d_1. \]

PROOF.

Recall \( n_{d_1+1} = v_{d_1+1} \). Assumption 4.2.5 then guarantees the first bound. Since \( \mu_s \) and all the \( n_i \) increase to infinity and all the \( \sigma_{0i} \) are finite, the second two terms are either bounded or unbounded together. Now by Assumption 4.2.3, each \( v_i = n_i^2 \) is the same order of magnitude as \( n_i \). But \( \mu_s = \dim \mathcal{K}_s \) defined in Section A.1, thus
\[ \hat{m}_s = \dim(L(\tilde{U}_{i_s} \ldots \tilde{U}_{p_1})) - \dim(L(\tilde{U}_{i_{s+1}} \ldots \tilde{U}_{p_1})). \]

Since

\[ \dim(L(\tilde{U}_{i_s} \ldots \tilde{U}_{p_1})) = m_i \text{ for any } i \in S_s \text{ and } \dim(L(\tilde{U}_{i_s} \ldots \tilde{U}_{p_1})) \leq \Sigma m_j, \]

it follows that \[ m_i - \Sigma_{j \in S_{s+1}^*} m_j \leq \hat{m}_s \leq m_i - \Sigma_{j \in S_s^*} m_j \] for any \( i \in S_s \). But then

\[ \frac{\hat{m}_s}{m_i} \text{ is bounded by } \Sigma_{j \in S_s} m_j - m_i - \Sigma_{j \in S_{s+1}^*} m_j \]

\[ \text{and} \quad \frac{\hat{m}_s}{m_i} \to \Sigma_{j \in S_s} \rho_{ji} < \infty \text{ and } \frac{\hat{m}_s}{m_i} \to 1 - \Sigma_{j \in S_{s+1}^*} \rho_{ji} = 1 \]

because \( \rho_{ji} = 0 \) for \( i \in S_s, j \in S_{s+1}^* \). Thus each \( \hat{m}_s \) has the same order of magnitude as each \( m_i \) for \( i \in S_s \). It is thus sufficient to show that

\[ \min_{i=0,1,\ldots,i_{s+1}-1} n_i = n_j \]

where \( j \in S_s \) at least for \( n \) large. However for \( t < s \) and \( i \in S_s \) and \( j \in S_t \), \( \rho_{ji} = \lim_{n \to \infty} \frac{m_j}{m_i} = + \infty \) by definition of the sets \( S_s \). Therefore, the minimum \( m_i \) and hence the minimum \( n_i \) must eventually equal \( n_j \) for some \( j \in S_s \). But then the minimum \( n_i \) and \( \hat{m}_s \) have the same order of magnitude and the second and third expressions are both bounded.

For the last two statements assume without loss of generality, that \( i \geq j \geq k \). Then \( i \in S_s, j \in S_t \) with \( s \geq t \). If \( s=t \), either \( m_i \) or \( m_j \) could be the minimum but then \( n_i \) and \( n_j \) both have the same order of magnitude and \( n_i n_j \) has the same order of magnitude as
\[
\min[m_i, m_j] \leq \min[m_i, m_j, m_k]. \text{ Hence the first expression is bounded and } \\
\text{since } n_k \to \infty \text{ the second converges to zero. If } s > t \text{ then } m_i \text{ will } \\
eventually be the minimum and } n_j \text{ and } n_k \text{ are both of greater order of} \\
magnitude than } n_i. \text{ Thus both expressions converge to zero.} \\

\text{Since all expressions are bounded, a common bound for all of} \\
\text{them can be chosen.} \quad \|

\text{A.4. Lemmas Used to Prove Theorem 4.4.1} \\

This section contains the lemmas required to prove Theorem 4.4.1. \\
Each lemma is stated and proved in a separate subsection. These \\
lemmas are referred to in the proof of Theorem 4.4.1 given in Section \\
4.5.
A.4.1. Proof of Conclusion 4.4.1.i—the Positive Definiteness of $\tilde{J}$

**Lemma A.4.1.** The $p \times p$ matrix $\tilde{J}$ defined by

$$
[J]_{ij} = \lim_{n \to \infty} \left[-\delta_0 \left( \frac{\delta^2 \chi_j^2}{\delta^2 \phi_1 \delta^2 \phi_j} \right) \right]_{\tilde{\lambda}_n = \tilde{\lambda}_{0n}}, \quad i,j = 1,2,\ldots,p\text{ is positive definite.}
$$

**Proof.**

It was shown in Section 4.3 that $J = \begin{bmatrix} \tilde{C}_0 & 0 \\ 0 & \tilde{C}_1 \end{bmatrix}$ and that $\tilde{C}_0$ was positive definite by Assumption 4.2.5. It remains to show that the $(p_1+1) \times (p_1+1)$ matrix $\tilde{C}_1$ is positive definite, where $[\tilde{C}_1]_{ij} = \frac{1}{2} \lim_{n \to \infty} \frac{1}{n_i n_j} \text{tr} \left( T^{-1}_0 \tilde{G}_i T^{-1}_0 \tilde{G}_j \right), \quad i,j = 0,1,\ldots,p_1.$

Let $b_0, b_1, \ldots, b_{p_1}$ be arbitrary constants, not all zero. It is required to show that

$$
P_1 P_1 \sum_{i=0}^{p_1} \sum_{j=0}^{p_1} b_i b_j \tilde{C}_1_{ij} > 0.
$$

But

$$
2 \sum_{i=0}^{p_1} \sum_{j=0}^{p_1} b_i b_j (\tilde{C}_1)_{ij} = P_1 P_1 \sum_{i=0}^{p_1} \sum_{j=0}^{p_1} b_i b_j \sum_{n_i n_j} \frac{1}{n_i n_j} \text{tr} \left( T^{-1}_0 \tilde{G}_i T^{-1}_0 \tilde{G}_j \right),
$$

$$
= \lim_{n \to \infty} \text{tr} \left( \sum_{i=0}^{p_1} \sum_{j=0}^{p_1} \frac{b_i \tilde{C}_i}{n_i} \right) \text{tr}^{-1}_0 \left( \sum_{j=0}^{p_1} \sum_{j=0}^{p_1} \frac{b_j \tilde{C}_j}{n_j} \right)
$$
because finite sums and limits and tracing interchange,

\[ = \lim_{n \to \infty} \text{tr} \left[ T^{-1}_0 \left( \sum_{i=0}^{p_1} \frac{1}{n_i} \tilde{G}_i \right) \right]^2 \]

\[ = \lim_{n \to \infty} \sum_{k=1}^{n} \lambda_k^2 \left[ T^{-1}_0 \left( \sum_{i=0}^{p_1} \frac{1}{n_i} \tilde{G}_i \right) \right]. \]

Thus the object of attention is positive for any finite \( n \) and the only problem is that it might degenerate in the limit. The proof that the limit is indeed positive proceeds as follows.

Suppose \( b_0 \neq 0 \); then some of the characteristic roots of

\[ T^{-1}_0 \left( \sum_{i=0}^{p_1} \frac{1}{n_i} \tilde{G}_i \right) \]

can be identified. Without loss of generality, write

\[ T_0 = \sum_{i=0}^{p_1} \sigma_i \tilde{G}_i \]

for the rest of this proof. To find some of the characteristic vectors in this case note that there is a space of dimension \( n_0^2 \) orthogonal to \( \mathcal{L}(\tilde{U}_1, \ldots, \tilde{U}_{p_1}) \). (See Section 4.2.)

Let an orthonormal basis for this space be \( \tilde{H}_0 \). Then \( U_0' \tilde{H}_0 = 0 \) and hence \( \tilde{G}_i \tilde{H}_0 = 0 \) for \( i=1,2,\ldots,p_1 \). This yields
\[
\left( \sum_{i=0}^{p_1} \frac{b_i}{n_i} G_i \right) \mathbb{H}_0 = \frac{b_0}{n_0} \mathbb{I} \cdot \mathbb{H}_0
\]

since \( G_0 = \mathbb{I} \),

\[
= \frac{b_0}{\sigma_{00}n_0} \left( \sigma_{00} \mathbb{I} \right) \mathbb{H}_0
\]

\[
= \frac{b_0}{\sigma_{00}n_0} \left( \sum_{i=0}^{p_1} \sigma_{0i} G_i \right) \mathbb{H}_0
\]

\[
= \frac{b_0}{\sigma_{00}n_0} \Sigma \mathbb{H}_0.
\]

Thus \( \Sigma^{-1} \left( \sum_{i=0}^{p_1} \frac{b_i}{n_i} G_i \right) \mathbb{H}_0 = \frac{b_0}{\sigma_{00}n_0} \mathbb{H}_0 \); then there are \( \nu_0(=n_0^2) \) characteristic vectors whose characteristic roots are \( \frac{b_0}{\sigma_{00}n_0} \). Then

\[
\text{tr} \left[ \Sigma^{-1} \left( \sum_{i=0}^{p_1} \frac{b_i}{n_i} G_i \right) \right]^2 \geq n_0^2 \left[ \frac{b_0}{\sigma_{00}n_0} \right]^2
\]

\[
= \frac{b_0^2}{\sigma_{00}^2} > 0
\]

Now let \( b_j = 0 \) for all \( j \notin S^* \); i.e. let \( s \) be the smallest index such that \( b_i \neq 0 \) for some \( i \in S_s^* \). Now observe \( \frac{|b_i|}{n_i} \) for \( i \in S_s^* \).
of these must be the largest in the sense that \( \lim_{n \to \infty} \frac{n_i}{|b_i|} \leq 1 \) for \( j \in S_i, j \neq i \). Consider this \( i \) fixed and now consider vectors belonging to \( \mathcal{E}(U_i) \). Also without loss of generality let \( b_i > 0 \) (clearly \( b_i \neq 0 \)). Now use Lemma B.5 to show that a number of characteristic roots of \( \sum_{j=0}^{\infty} \left( \sum_{j=0}^{n_j} \frac{b_j}{n_j} G_j \right) \) have a lower bound of the proper order of magnitude; that is, consider

\[
\inf_{\bar{x} \neq 0} \frac{x'}{\bar{x}} \left( \sum_{j=0}^{n_j} \frac{n_j}{b_j} G_j \right) \bar{x}
\]

where \( \bar{x} \) is a subspace of \( \mathcal{E}(U_i) \). Equivalently, consider

\[
\inf_{\gamma' \neq 0} \frac{\gamma'}{\gamma} \left( \sum_{j=0}^{n_j} \frac{n_j}{b_j} G_j \right) \gamma
\]

where restrictions are placed on possible \( \gamma \) vectors to restrict consideration to the appropriate subspace. \( \mathcal{E}(U_i) \) is a space of dimension \( m_i \); the number of restrictions placed on \( \gamma \) will determine how many characteristic roots there are greater than the lower bound which is eventually arrived at.
First, note that for any \( j \in S^*_{s+1} \) (recall \( i \in S_s \)) the matrix \( U^U_{j,i} \)

is an \( m_j \times m_i \) matrix and hence has rank at most \( m_j \) (because \( m_j \) has

smaller order of magnitude than \( m_i \)). Hence by restricting \( \chi \) to an

\( m_i - m_j \)-dimensional space it can be insured that \( U^U_{j,i} \chi = 0 \). Thus by

restricting \( \chi \) to a space of at worst (i.e. smallest) dimension \( m_i - \Sigma \)

\( j \in S^*_{s+1} \)

it can be insured that \( U^U_{j,i} \chi = 0 \) and hence that \( U^U_{j,i} \chi = 0 \)

\( j \in S^*_{s+1} \)

for \( j \in S^*_{s+1} \). This restricts consideration to

\[
\inf_{\chi \neq 0} \frac{\chi^T U^U_{j,i} (\Sigma_{j \in S^*_{s+1}} \frac{b_j}{n_j} G_j) U^U_{j,i} \chi}{\chi^T U^U_{j,i} (\Sigma_{j \notin S^*_{s+1}} \sigma_{0j} G_j) U^U_{j,i} \chi},
\]

\( \chi \) restricted

because \( b_j = 0 \) for \( j \notin S^*_{s} \). Now note that \( \chi^T U^U_{j,i} \chi = \chi^T U^U_{j,i} \chi = \chi^T U^U_{j,i} \chi > 0 \)

because \( D^U_j \) is a nonsingular diagonal matrix by the definition of \( U^U_j \).

Thus the expression under consideration can be rewritten as follows:

\[
\inf_{\chi \neq 0} \frac{\chi^T U^U_{j,i} \chi + \Sigma_{j \in S^*_{s+1}} \frac{b_j}{n_j} \chi^T G_j U^U_{j,i} \chi}{\chi^T U^U_{j,i} \chi + \Sigma_{j \notin S^*_{s+1}} \sigma_{0j} U^U_{j,i} \chi + \sigma_{00} \chi^T U^U_{j,i} \chi + \sigma_{01} \chi^T U^U_{j,i} \chi}
\]

\( \chi \) restricted

\( j \notin S^*_{s+1} \)

\( j \neq 0 \)
\[
\begin{align*}
\chi \text{ restricted} & \quad \frac{1 + \sum_{j \in S_s} \frac{b_j n_j}{n_j b_1} \frac{\chi' U_j U_i U_i' U_i' \chi}{\chi D_j^2 \chi}}{\chi D_1 \chi} \\
& = \inf_{\chi \geq 0} \frac{b_i}{n_i} \cdot \\
\chi \text{ restricted} & \quad \sum_{j \in S_s^*} \sigma_{0j} \frac{\chi' U_j U_i U_i' U_i' \chi}{\chi D_j^2 \chi} + \sigma_{00} \frac{\chi D_j \chi}{\chi D_1 \chi} + \sigma_{01}
\end{align*}
\]

by dividing each term by \(\chi' D_1 \chi\) and factoring out \(\frac{b_i}{n_i}\),

\[
\begin{align*}
\chi \text{ restricted} & \quad \frac{1 + \sum_{j \in S_s} \frac{b_j n_j}{n_j b_1} \frac{\chi' D_j D_1^{-1} U_j U_i U_i' U_i' D_j^{-1} \chi}{\chi D_1 \chi}}{\chi D_1 \chi} \\
& = \inf_{\chi \geq 0} \frac{b_i}{n_i} \cdot \\
\chi \text{ restricted} & \quad \sum_{j \in S_s^*} \sigma_{0j} \frac{\chi' D_j D_1^{-1} U_j U_i U_i' U_i' D_j^{-1} \chi}{\chi D_1 \chi} + \sigma_{00} \frac{\chi' D_j D_1^{-1} \chi}{\chi D_1 \chi} + \sigma_{01}
\end{align*}
\]

\[
\begin{align*}
\xi \text{ restricted} & \quad \frac{1 + \sum_{j \in S_s} \frac{b_j n_j}{n_j b_1} \frac{\xi' D_j D_1^{-1} U_j U_i U_i' U_i' D_j^{-1} \xi}{\xi D_1 \xi}}{\xi D_1 \xi} \\
& = \inf_{\xi \geq 0} \frac{b_i}{n_i} \cdot \\
\xi \text{ restricted} & \quad \sum_{j \in S_s^*} \sigma_{0j} \frac{\xi' D_j D_1^{-1} U_j U_i U_i' U_i' D_j^{-1} \xi}{\xi D_1 \xi} + \sigma_{00} \frac{\xi' D_j D_1^{-1} \xi}{\xi D_1 \xi} + \sigma_{01}
\end{align*}
\]

where \(\xi = D_1 \chi\),
because $\lambda_{\text{max}}(D^{-1}_d) \leq 1$ since each diagonal element of $D_d$ is greater than 1. Thus matrices of the form $D^{-1}_d U'_j U'_j D^{-1}_d$ must be studied. Consider the trace of such a matrix.

$$\text{tr}(D^{-1}_d U'_j U'_j D^{-1}_d) = \text{tr}[(U'_j U'_j D^{-1}_d)' (U'_j U'_j D^{-1}_d)]$$

$$= \sum_{k=1}^{m_i} \sum_{l=1}^{m_j} (U'_j U'_j D^{-1}_d)_{lk}^2,$$

because $\text{tr} A^T A = \sum \sum a^2_{lk} \text{ for any matrix } A$. But now let the columns of $U_j$ be $U^{(j)}_1, \ldots, U^{(j)}_{m_j}$; then since $D_d = U'_j U'_j$,

$$[U'_j U'_j D^{-1}_d]_{lk} = \frac{U^{(j)}_l U^{(i)}_k}{U^{(i)}_k U^{(i)}_k} \cdot$$
Observe that \( \frac{u_{(j)'} - u_{(i)}}{u_{(i)} - u_{(i)}} \leq 1 \) because all columns contain only zeros and ones and hence the numerator counts matches of ones in \( u_{(j)} \) and \( u_{(i)} \) and the denominator counts the ones in \( u_{(i)} \). Since there can be no more matches than there are ones in \( u_{(i)} \) the inequality is true.

Furthermore

\[
\sum_{j=1}^{m_j} \frac{u_{(j)'} - u_{(i)}}{u_{(i)} - u_{(i)}} = 1
\]

because \( \sum_{j=1}^{m_j} u_{(j)} \) is a \( n \times 1 \) vector of ones by definition of the \( U_j \).

Hence

\[
\sum_{j=1}^{m_j} \left( \frac{u_{(j)'} - u_{(i)}}{u_{(i)} - u_{(i)}} \right)^2
\]

is a sum of squares of items all of which are between zero and one and which add up to one. Then the sum of squares has a maximum of one which occurs when one of the summands equals one and the others are zero. Otherwise the sum will be less than one (usually much less). At this point Assumption 4.2.4 quite naturally applies. It says that for \( j \in S_s, j \neq i \) there exist constants \( R_1 \) and \( R_2 \) such that except for \( R_{m_i} \) of the \( u_{(i)} \), the quantities

\[
\sum_{j=1}^{m_j} \left( \frac{u_{(j)'} - u_{(i)}}{u_{(i)} - u_{(i)}} \right)^2
\]

are less than \( R_2 \). Thus
\[
\text{tr}(D^{-1}_1 U'_1 U U'_2 U D^{-1}_2) = \sum_{k=1}^{m_1} \sum_{\ell=1}^{m_j} \left( \frac{u'_\ell(j)^{\prime}}{\bar{u}_k(i)} \frac{u_k(i)}{\bar{u}_k(i)} \right)^2 \\
\leq (R_1 m_1) l + (m_1 - R_1 m_1) R_2 \\
= \begin{cases} \\
m_1 \left[ R_1 + (1 - R_1) R_2 \right] \\
\end{cases} \\
\leq \frac{m_1}{N(S_s) + 1},
\]

where \( N(S_s) \) is the number of indices in \( S_s \) by Assumption 4.2.4. Of course, for \( j \notin S_s \) the bound \( \text{tr}(D^{-1}_1 U'_1 U U'_2 U D^{-1}_2) \leq m_1 \) still holds, since

\[
\sum_{\ell=1}^{m_j} \left( \frac{u'_\ell(j)^{\prime}}{\bar{u}_k(i)} \frac{u_k(i)}{\bar{u}_k(i)} \right)^2 \leq 1, \quad k=1,2,\ldots,m_1.
\]

Now if \( \tilde{A} \) is any positive semidefinite \( m \times m \) matrix with \( \text{tr}(\tilde{A}) \leq K_1 m \) it is clear that at most \( \frac{K_1}{K_2} m \) of the characteristic roots of \( \tilde{A} \) can be greater than \( K_2 \). (If they were, \( \text{tr} \ A > K_2 \cdot \frac{K_1}{K_2} m = K_1 m \), a contradiction.) This fact with \( K_2 = 2p_1(N(S_s) + 1) \) and \( K_1 = 1 \) implies that the denominator of \((**): \Sigma \left[ \text{tr} \left( S^*_s + \sigma^{0+} \sigma^{0+} \sigma^{0i} \right) \right] \) is less than \( 2p_1(N(S_s) + 1) \).

\[
\sum_{j \in S_s^*} \left[ \text{tr} \left( S^{0i} + \sigma^{00+} \sigma^{0i} \right) \right]_{j \neq i} \left[ \sum_{j \neq 0} \right]
\]
is restricted so that it contains no part of the \( \frac{K_1}{K_2} m_1 \) characteristic
vectors that go along with the "offending" characteristic roots.
Furthermore the numerator of (***) is equal to

\[
1 - \sum_{j \in S_s} b_{ij} b_{n_j}^{-1} \frac{\varepsilon^{D_{-1} U_{-1} (\sum_{j \in S_s} U_j U_j') U_{-1} D_{-1}^{-1} \varepsilon}}{\varepsilon^j \varepsilon^i}
\]

\[\geq 1 - \left(1 + \frac{1}{2N(S_s)} \right) \sum_{j \in S_s} \frac{\varepsilon^{D_{-1} U_{-1} (\sum_{j \in S_s} U_j U_j') U_{-1} D_{-1}^{-1} \varepsilon}}{\varepsilon^j \varepsilon^i}
\]

since \( \left| \frac{b_{ij} b_{n_j}^{-1}}{b_{n_j} b_{ij}} \right| \leq 1 + \frac{1}{2N(S_s)} \) for all \( j \in S_s \) \( j \neq i \) for all \( n > \) beyond some
point as \( n \to \infty \) because \( \lim_{n \to \infty} \left| \frac{b_{ij} b_{n_j}^{-1}}{b_{n_j} b_{ij}} \right| \leq 1 \) by choice of \( i \),

\[
\frac{\varepsilon^{D_{-1} U_{-1} (\sum_{j \in S_s} U_j U_j') U_{-1} D_{-1}^{-1} \varepsilon}}{\varepsilon^j \varepsilon^i}
\]

\[= 1 - \left(1 + \frac{1}{2N(S_s)} \right) \sum_{j \in S_s} \frac{\varepsilon^{D_{-1} U_{-1} (\sum_{j \in S_s} U_j U_j') U_{-1} D_{-1}^{-1} \varepsilon}}{\varepsilon^j \varepsilon^i}
\]

But now
\[
\text{tr}[D_{-1} U_{-1} (\sum_{j \in S_s} U_j U_j') U_{-1} D_{-1}^{-1}] \leq \frac{N(S_s) - 1}{N(S_s) + 1} \cdot m_1
\]
since the trace of a sum is the sum of the traces and since each individual trace is bounded above and there are $N(S_s)-1$ traces. Now using the above argument about traces and characteristic roots again with $K_2 = 1 - \frac{1}{2N(S_s)}$ and $K_1 = \frac{N(S_s)}{N(S_s)} + \frac{1}{2}$, if $\xi$ is further restricted so that it contains no part of the characteristic vectors associated with characteristic roots of $\sum_{j \in S_s} \sum_{j \neq i} (\Sigma_{j \in S_s} U_j U'_j) U_{j \neq i} D_j^{-1}$ which are greater than $1 - \frac{1}{2N(S_s)}$ then the numerator of (**) is greater than

$$1 - \left(1 + \frac{1}{2N(S_s)}\right) \left(1 - \frac{1}{2N(S_s)}\right) = \frac{1}{4N(S_s)^2} > 0.$$  

It is now necessary to count the number of restrictions which have been made on $\xi$ at this point. At most $\sum_{j \in S^*} m_j$ were made for $j \neq S + 1$

the first restrictions (when $\xi$ was $\psi$); at most

$$\sum_{j \neq i} K_2 \frac{m_i}{K_2} = \sum_{j \neq i} \frac{m_i}{2p_1(N(S_s)+1)} \leq \frac{m_i}{2(N(S_s)+1)}$$ (because there are at most $p_1$ terms in the summation) more restrictions were made to bound the denominator; at most $\frac{K_1}{K_2} m_i = \frac{(N(S_s)-1)m_i}{(N(S_s)+1)(1-\frac{1}{2N(S_s)})}$. More restrictions were
made to bound the numerator. Thus the total number of restrictions is
less than or equal to the sum of these numbers and the dimension of the
linear space over which the inf may be taken will be greater than or
equal to \(m_i\) minus this sum. In fact this dimension is greater than
or equal to

\[
\frac{m_i}{2(N(S_s)+1)} - \frac{m_i(N(S_s)-1)}{(N(S_s)+1)(1-\frac{1}{2N(S_s)})} - \sum_{j \in S_{s+1}^*} m_j
\]

\[
= m_i \left[ 1 - \frac{1}{2(N(S_s)+1)} - \frac{N(S_s)-1}{(N(S_s)+1)(1-\frac{1}{2N(S_s)})} - \sum_{j \in S_{s+1}^*} \left( \frac{m_j}{m_i} \right) \right]
\]

\[
= m_i \left[ 2 - \frac{1}{N(S_s)} \right] \frac{1}{2(N(S_s)+1)(1-\frac{1}{2N(S_s)})} - \sum_{j \in S_{s+1}^*} \left( \frac{m_j}{m_i} \right) \right]
\]

But \(\frac{m_j}{m_i} \to 0\) as \(n \to \infty\) for each \(j \in S_{s+1}^*\) so that for \(n\) large enough the

last expression in brackets is greater than \(K_3= \frac{1}{4} \left[ \frac{2 - \frac{1}{N(S_s)}}{(N(S_s)+1)(1-\frac{1}{2N(S_s)})} \right]\),

which is positive.
Now combine the bounds obtained on the numerator and denominator of (***) to obtain

\[
\inf_{\|x\|=0} \frac{b_i}{n_i} \geq \frac{1}{\frac{4(N(S_s)^2)}{2p_1(N(S_s)+1) \sum_{j \in S^*_s \setminus j=1}^{s+1} \sigma_{0j} + \sigma_{00} + \sigma_{01}}} \geq \frac{1}{\frac{p_1 b_i}{n_i}} K_4 > 0.
\]

Now Lemma B.5 states that there are at least \(K_3 m_1\) characteristic roots of \(\Sigma^{-1}( \sum_{j=0}^{n_j} \Sigma_{j} \sim 0)\) which are greater than or equal to \(\frac{b_i}{n_i} K_4\).
This yields

\[
\text{tr}\left[ (\Sigma_0^{-1} \left( \sum_{j=0}^{p_1} b_{i,j} \right) \right]^2 \geq K_3 m_1 \frac{b_1^2}{n_i} K_4^2
\]

\[
= \frac{m_i}{v_i} \frac{b_1^2}{v_i} K_3 K_4^2
\]

\[
\sim \frac{1}{n_i} \frac{b_1^2}{v_i} K_3 K_4^2 > 0
\]

Thus for all choices of \( b \), \( b^T C_{b} b > 0 \) and hence \( C_{b} \) is positive definite.

This proves that \( \tilde{J} \) is positive definite and concludes the proof of Lemma A.4.1.
A.4.2. Verification of Condition 3.3.1.ii—Asymptotic Normality of

\[ \frac{\partial \lambda}{\partial \psi} \bigg| _{\psi = \psi_{On}}. \]

**Lemma A.4.2.** For $\psi$, $\lambda(\psi, \varphi)$, and $J$ as defined in Section 4.5,

\[ \frac{\partial \lambda(\psi, \varphi)}{\partial \varphi} \bigg| _{\varphi = \varphi_{On}} \sim \mathcal{N}(0, J). \]

**Proof.**

Let the vector \( \frac{\partial \lambda}{\partial \psi} \bigg| _{\psi = \psi_{On}} = a_{n} = \begin{pmatrix} a_{n}^{(0)} \\ a_{n}^{(1)} \end{pmatrix} \), where \( a_{n}^{(0)} \) is a \( p_0 \times 1 \) vector defined by

\[ a_{n}^{(0)} = \frac{\partial \lambda}{\partial \varphi} \bigg| _{\varphi = \varphi_{On}} = \frac{1}{n_{p_{1} + 1}} \chi_{p_1 + 1}^{-1} (\chi - \chi \frac{\tilde{v}_{0}}{n_{p_{1} + 1}}), \]

and \( a_{n}^{(1)} \) is a \( (p_{1} + 1) \times 1 \) vector defined by

\[ [a_{n}^{(1)}]_{i} = \frac{\partial \lambda}{\partial \sigma_{i}} \bigg| _{\psi = \psi_{On}}. \]
\[
\frac{1}{2n_i} \left[ -\text{tr} T_0^{-1} G_1 + (\chi - \frac{\varepsilon_0}{n_{p_1} + 1}) T_0^{-1} G_1 T_0^{-1} (\chi - \frac{\varepsilon_0}{n_{p_1} + 1}) \right], \\
i = 0, 1, \ldots, p_1.
\]

But
\[
T_0 = \sum_{i=0}^{p_1} \frac{[T_0 n_i^j]}{n_i} G_1
\]

\[
P_1 = \sum_{i=0}^{p_1} \sigma_{0i} G_1
\]

\[
= \Sigma_0 ,
\]

and
\[
\frac{\varepsilon_0}{n_{p_1} + 1} = \frac{\varepsilon_0 n_i}{n_{p_1} + 1}
\]

\[
= \varepsilon_0 ,
\]

for all \( n \), so in the remainder of the proof, those substitutions will
be used. Since \( \Sigma_0 \) is positive definite, let \( \Sigma_0 = AA' \), which implies
\( \Sigma_0^{-1} = A^{-1} A^{-1} \). (Recall \( A^{-1} = (A')^{-1} = (A^{-1})' \).) Let \( z \) be an \( n \times 1 \) vector
defined by \( z = A^{-1}(x - x_0) \); then \( z \sim \mathcal{N}(0, \Sigma_0) \). \( z_n^{(0)} \) and \( z_n^{(1)} \) are re-defined in terms of \( z \) as follows:
\[ a_n^{(0)} = \frac{1}{n^{p_1+1}} X'_A^{-t} z, \]

\[ [\hat{a}_n^{(0)}]_i = \frac{1}{2n_1} (z A^{-1} G_4 A^{-t} z - \text{tr} \Sigma_0^{-1} G_4), \quad i=0,1,\ldots,p_1. \]

Let \( \delta_{p1} = \left( \begin{array}{c} \delta^{(0)} \\ \delta^{(1)} \end{array} \right) \), partitioned the same as \( \hat{a}_n \). It is then sufficient to show that for any choice of \( \delta \), \( \delta a_n \) is asymptotically normal with mean zero and variance \( \delta' J \delta \).

Let

\[ W(\delta, n) = \delta' a_n = \sum_{i=0}^{p_1} \frac{\delta^{(1)}_i}{2n_1} [z' (A^{-1} G_4 A^{-t}) z - \text{tr} \Sigma_0^{-1} G_4] + \frac{\delta^{(0)}'}{n^{p_1+1}} X'_A^{-t} z \]

\[ = z' p(\delta, n) z - \text{tr} \Sigma(\delta, n) + f'(\delta, n) z, \]

where \( P(\delta, n) = A^{-1} \left( \sum_{i=0}^{p_1} \frac{\delta_i^{(1)}}{2n_1} G_4 \right) A^{-t} \) and \( f(\delta, n) = \frac{1}{n^{p_1+1}} A^{-1} \Sigma^{(0)} \).

Now calculate the characteristic function of \( W(\delta, n) \). For each \( n \) there exists an orthogonal matrix \( F(\delta, n) \) and a diagonal matrix \( \Lambda(\delta, n) \) \( (\Lambda(\delta, n), k = \lambda_k) \) such that \( F(\delta, n) = F(\delta, n) \Lambda(\delta, n) F(\delta, n) \). Of course, \( \Lambda \) contains the characteristic roots of \( F \) and \( F \) the characteristic vectors; the decomposition is possible because \( F \) is symmetric.
Then the $n \times 1$ vector $\mathbf{w}(n, \delta) = H(n, \delta) \sim \mathcal{N}(0, I_n)$ and
\[
\mathbb{W}(n, \delta) = \mathbb{w}(n, \delta) A(n, \delta) \mathbb{w}(n, \delta) - \text{tr} A(n, \delta) + g(n, \delta) \mathbb{w}(n, \delta),
\]
where
\[
g(n, \delta) = \mathbb{w}(n, \delta) F(n, \delta) \text{ and tr } F(n, \delta) = \text{tr} A(n, \delta).
\]
Now the dependence on $\delta$ and $n$ is suppressed in the notation, yielding $\mathbb{W} = \mathbb{w}' A \mathbb{w} - \text{tr} A + g \mathbb{w}$, where $\mathbb{w} \sim \mathcal{N}(0, I_n)$. The characteristic function of $\mathbb{W}$, $\phi_w(t)$ is given, for any $t$, by
\[
\phi_w(t) = \mathcal{S}_0 \{ e^{it\mathbb{W}} \}
= \mathcal{S}_0 \{ e^{it(\mathbb{w}' A \mathbb{w} - \text{tr} A + g \mathbb{w})} \}
= e^{-it \text{tr} A} \mathcal{S}_0 \{ e^{it g' \mathbb{w} + it \mathbb{w}' A \mathbb{w}} \}
\]
Thus Lemma B.2 applies, yielding
\[
\phi_w(t) = e^{-it \text{tr} A} \log |I - 2it A|^{\frac{1}{2}} \ e^{-\frac{1}{2} t^2 g' (I - 2it A)^{-1} g}.
\]
Then $\log \phi_w(t) = -it \text{tr} A \frac{1}{2} \log |I - 2it A|^{\frac{1}{2}} \ e^{-\frac{1}{2} t^2 g' (I - 2it A)^{-1} g}$

\[
= -it \sum_{k=1}^{n} \lambda_k - \frac{1}{2} \sum_{k=1}^{n} \log(1 - 2it \lambda_k) - \frac{1}{2} t^2 \sum_{j=0}^{\infty} \binom{2j}{j} g_j,
\]
where the last expansion is valid so long as $|\lambda_k| < \frac{1}{2t}$, which is true for $n$ sufficiently large. Note that
\[ \lambda_k = \lambda_k(F) = \lambda_k(A^{-t} \sum_{i=0}^{p_1} \frac{\delta^{(1)}_i}{2n_1} G_i A^{-1}) \]

by Lemma B.9,

\[ = \lambda_k(A^{-1} A^{-t} \sum_{i=0}^{p_1} \frac{\delta^{(1)}_i}{2n_1} G_i) \]

\[ = \lambda_k(\sum_{i=0}^{p_1} \frac{\delta^{(1)}_i}{2n_1} G_i x_k' G_i x_k) \]

by Lemma B.6, where \( x_k \) is the associated characteristic vector.

Therefore

\[ |\lambda_k| \leq \frac{p_1 \sum_{i=0}^{p_1} \frac{\delta^{(1)}_i}{2n_1} |x_k' G_i x_k|}{\sum_{i=0}^{p_1} \sigma_{0i} x_k' G_i x_k} \]

\[ \leq \max_{i=0,1,\ldots,p_1} \frac{|\delta^{(1)}_i|}{2n_1 \sigma_{0i}} \]

by Lemma B.3. The last quantity converges to zero since \( n_1 \to \infty \) for all \( i \); this certainly can be made less than \( \frac{1}{2t} \) for \( n \) large enough.
Continuing the expansion of \( \log \phi_W(t) \),

\[
\log \phi_W(t) = -it \sum_{k=1}^{n} \lambda_k - \frac{1}{2} t \sum_{k=1}^{n} \log(1-2it \lambda_k) - \frac{1}{2} t^2 \sum_{j=0}^{\infty} (2it \lambda_k)^j g.
\]

The second term is

\[
-\frac{1}{2} \sum_{k=1}^{n} \log(1-2it \lambda_k) = \frac{1}{2} \sum_{k=1}^{n} \sum_{j=1}^{\infty} (2it \lambda_k)^j \cdot \frac{1}{j}
\]

\[
= \frac{1}{2} \sum_{k=1}^{n} [2it \lambda_k - 2t^2 \lambda_k^2 + \sum_{j=3}^{\infty} (2it \lambda_k)^j \cdot \frac{1}{j}],
\]

while the last term is

\[
-\frac{1}{2} t^2 \sum_{j=0}^{\infty} (2it \lambda_k) g_j = -\frac{1}{2} t^2 g' g - \frac{1}{2} t^2 \sum_{j=1}^{\infty} g_j' (2it \lambda_k) g.
\]

Combining terms,

\[
\log \phi_W(t) = -it \sum_{k=1}^{n} \lambda_k + it \sum_{k=1}^{n} \lambda_k - \frac{t}{2} \sum_{k=1}^{n} 2 \lambda_k^2 - \frac{t}{2} g' g
\]

\[
+ \frac{1}{2} \sum_{k=1}^{n} \sum_{j=3}^{\infty} (2it \lambda_k)^j \cdot \frac{1}{j}
\]

\[
-\frac{1}{2} t^2 \sum_{j=1}^{\infty} (2it \lambda_k)^j g_j' A^j g.
\]
But recall that

\[ g' g = f' f' \]

\[ = f' f \]

\[ \sim \]

\[ = \frac{1}{n_{p_1 + 1}} \delta^{(0)} (X') \chi_{A^{-t}} \sim \zeta_{A^{-1}} \sim (0) \]

\[ = \frac{1}{\nu_{P_1 + 1}} \delta^{(0)} (X') \chi_{\zeta_{0}^{-1}} \chi_{\zeta_{0}^{(0)}} \]

\[ \rightarrow \delta^{(0)}, \zeta_{0}^{(0)} \]

by Assumption 4.2.5. Furthermore

\[ 2 \sum_{k=1}^{n} \lambda_{k}^{2} = 2 \text{tr} A^{2} = 2 \text{tr} F^{2} \]

\[ = 2 \text{tr} (A^{-1} \sum_{i=0}^{p_1} \frac{1}{2n_i} \delta^{(1)}_{i} A^{-1} + \sum_{j=0}^{p_1} \frac{1}{2n_j} \delta^{(1)}_{j} A^{-t}) \]

\[ = \sum_{i=0}^{p_1} \sum_{j=0}^{p_1} \delta^{(1)}_{i} \delta^{(1)}_{j} \cdot \frac{1}{2n_i n_j} \text{tr} \zeta_{A^{-1}} G_{i} A^{-1} G_{j} A^{-t} \]

\[ = \sum_{i=0}^{p_1} \sum_{j=0}^{p_1} \delta^{(1)}_{i} \delta^{(1)}_{j} \cdot \frac{1}{2n_i n_j} \text{tr} \zeta_{0}^{-1} G_{i} \zeta_{0}^{-1} G_{j} \]
\[
- \sum_{i=0}^{p_1} \sum_{j=0}^{p_1} \delta_i^{(1)} \delta_j^{(1)} (c_{ij})_{i,j}
= \delta^{(1)}_{\sim}, \quad c_{\sim} \delta^{(1)}_{\sim}
\]

Thus since \( J = \begin{pmatrix} c_0 & 0 \\ 0 & c_{\sim} \end{pmatrix} \) the first remaining terms of \( \log \phi_w(t) \) converge to \( \frac{1}{2} t^{2} \delta' J \delta \) which is what is required for normality. It remains to show the last two terms converge to zero.

Now
\[
S_1 = \left| \sum_{k=1}^{\infty} \sum_{j=3}^{\infty} (2it \lambda_k)^j \frac{1}{j} \right| \leq \sum_{k=1}^{\infty} \sum_{j=3}^{\infty} 2^{j} t^j |\lambda_k|^j.
\]

As previously noted \( \max |\lambda_k| \to 0 \) as \( n \to \infty \) so that \( 2t \max |\lambda_k| < \frac{1}{2} \) for \( n \) large enough and therefore
\[
S_1 \leq \sum_{k=1}^{n} \frac{2^{3} t^3 |\lambda_k|^3}{1-2t|\lambda_k|} \leq 16t^3 \sum_{k=1}^{n} |\lambda_k|^3 \leq 16t^3 \max |\lambda_k| \sum_{k=1}^{n} |\lambda_k|^2.
\]

But \( \sum_{k=1}^{n} |\lambda_k|^2 \to \frac{1}{2} \delta^{(1)}_{\sim}, \quad c_{\sim} \delta^{(1)}_{\sim} \leq \infty \) as previously noted and is therefore
bounded, and \(\max|\lambda_k| \to 0\) so that indeed \(S \to 0\).

\[
S_2 = \sum_{j=1}^{\infty} (2\pi t)^j \mathcal{A}_J \mathcal{A}_J^j \mathcal{A}_J^j
\]

\[
= \sum_{j=1}^{\infty} (2\pi t)^j \mathcal{A}_J \mathcal{A}_J^j \mathcal{A}_J^j \mathcal{F}_J^j
\]

\[
= \sum_{j=1}^{\infty} (2\pi t)^j \mathcal{A}_J \mathcal{A}_J^j \mathcal{F}_J^j \mathcal{F}_J^j
\]

\[
\leq \sum_{j=1}^{\infty} 2^j t \mathcal{A}_J \mathcal{A}_J^j \mathcal{F}_J^j \mathcal{F}_J^j
\]

\[
\leq \sum_{j=1}^{\infty} 2^j t \mathcal{A}_J \mathcal{A}_J^j \lambda_{\text{max}}(\mathcal{F}_J^j \mathcal{F}_J^j)
\]

by Lemma B.12. But \(\mathcal{F}_J \mathcal{F}_J \to \delta_0^{(0)}\), \(\mathcal{L}_0 \delta_0^{(0)} \to \infty\) and is therefore bounded and \(\mathcal{F}_J\) is symmetric so that

\[
\lambda_{\text{max}}(\mathcal{F}_J^j \mathcal{F}_J^j) = \lambda_{\text{max}}(\mathcal{F}_J^j)\]

\[
\leq \max_{k=1,2,\ldots,n} |\lambda_k(\mathcal{F}_J)|^j
\]

by Lemma B.14, so that

\[
\lambda_{\text{max}}^{\frac{1}{j}}(\mathcal{F}_J^j \mathcal{F}_J^j) \leq \max_{k=1}^{n} |\lambda_k(\mathcal{F}_J)|^j
\]

\[
= \max_{k=1}^{n} |\lambda_k|^j
\]
But again this converges to zero so that it may be assumed
\[ 2t \max |\lambda_k| < \frac{1}{2} \]
and therefore
\[ S_2 \leq \sum_{j=1}^{\infty} (2t \max |\lambda_k|)^j \]
\[ = \sum_{j=1}^{\infty} \frac{2t \max |\lambda_k|}{1-2t \max |\lambda_k|} \]
\[ \leq 4t \sum_{j=1}^{\infty} \max |\lambda_k| \]
\[ \to 0. \]

Thus \( \log \phi_W(t) \to -\frac{1}{2} t^2 \bar{\delta} J \bar{\xi} \) and since \( \bar{\mathcal{J}} \) is positive definite the limit is a legitimate characteristic function, continuous at \( t=0 \) and \( \bar{\mathcal{J}}=Q \). In fact \( e^{-\frac{1}{2} t^2 \bar{\delta} \bar{\mathcal{J}} \bar{\xi}} \) is the characteristic function of a random variable distributed as \( \mathcal{N}(0,\bar{\delta} \bar{\mathcal{J}} \bar{\xi}) \). This proves that \( W(\delta,\mathcal{N}) \overset{d}{=} \mathcal{N}(0,\bar{\delta} \bar{\mathcal{J}} \bar{\xi}) \) and since \( \delta \) is arbitrary that \( a_n \overset{d}{=} \mathcal{N}(0,\bar{\mathcal{J}}) \), which was to be proved.
A.4.3. Verification of Condition 3.3.1.iii--Convergence in Probability

\[ \frac{\frac{\partial^2 \lambda}{\partial \psi_i \partial \psi_j}}{\partial \psi_i} \bigg|_{\psi_i = \hat{\psi}_0} \to \text{Its Expected Value} \]

Lemma A.4.3. For \( \psi \) and \( \lambda(\xi, \psi) \) as defined in Section 4.5,

\[ \frac{\partial^2 \lambda(\xi, \psi)}{\partial \psi_i \partial \psi_j} \bigg|_{\psi = \hat{\psi}_0} \text{ converges in probability to its expected value,} \]

\[ i, j = 1, 2, \ldots, p. \]

PROOF:

It suffices to prove that the variance of each element of \( \frac{\partial^2 \lambda}{\partial \psi_i \partial \psi_j} \)
converges to zero as \( n \to \infty \). There are three forms of second derivatives, which are exhibited in Section 4.3. Each will be dealt with separately.

Since each element of \( \frac{\partial^2 \lambda}{\partial \beta i \beta j} \) has variance zero (being a constant),
the lemma is true for these derivatives.

To examine the derivatives of the form \( \frac{\partial^2 \lambda}{\partial \beta i \beta j} \), it is sufficient to show that for all \( p \times 1 \) vectors \( \xi \) such that \( \xi' \xi = 1 \), \( \text{Var}_0 \{ \phi_i(\xi) \} \to 0, i = 0, 1, \ldots, p_1 \), where
\[ \phi_1(\xi) = \mathbb{E} \left[ \frac{\frac{\partial^2 \lambda(\xi, \xi)}{\partial \xi_1^2}}{\xi_1 \mid \xi} \right] \]

\[ = \frac{1}{n_1^2 n_{p_1}^2} \mathbb{E} \left[ \frac{\partial^2 \lambda(\xi, \xi)}{\partial \xi_1^2} \right] \frac{\lambda_0}{n_{p_1}^2} \left( \lambda - \frac{\lambda_0}{n_{p_1}^2} \right) \]

From Section 4.3,

\[ = \frac{1}{n_1^2 n_{p_1}^2} \mathbb{E} \left[ \frac{\partial^2 \lambda(\xi, \xi)}{\partial \xi_1^2} \right] \frac{\lambda_0}{n_{p_1}^2} \left( \lambda - \lambda_0 \right) \]

Using the same substitutions as in Section A.4.2.

Then \( \delta_0[\phi_1(\xi)] = 0 \) and

\[ \text{Var}_0[\phi_1(\xi)] = \frac{1}{n_1^2 n_{p_1}^2} \mathbb{E} \left[ \frac{\partial^2 \lambda(\xi, \xi)}{\partial \xi_1^2} \right] \frac{\lambda_0}{n_{p_1}^2} \left( \lambda - \frac{\lambda_0}{n_{p_1}^2} \right) \]

\[ = \frac{1}{n_1^2 n_{p_1}^2} \mathbb{E} \left[ \frac{\partial^2 \lambda(\xi, \xi)}{\partial \xi_1^2} \right] \frac{\lambda_0}{n_{p_1}^2} \left( \lambda - \frac{\lambda_0}{n_{p_1}^2} \right) \]

\[ = \frac{1}{n_1^2 n_{p_1}^2} \mathbb{E} \left[ \frac{\partial^2 \lambda(\xi, \xi)}{\partial \xi_1^2} \right] \frac{\lambda_0}{n_{p_1}^2} \left( \lambda - \frac{\lambda_0}{n_{p_1}^2} \right) \]

(Recall that \( \Sigma_0 = AA' \).)

\[ \leq \frac{1}{n_{p_1}^2} \left[ \mathbb{E} \left[ \frac{\partial^2 \lambda(\xi, \xi)}{\partial \xi_1^2} \right] \right] \frac{1}{n_1^2} \lambda_{\text{max}} \left( A^{-1}G_1A^{-t} - A^{-1}G_1A^{-t} \right) \]
by definition of characteristic root,

\[
= \frac{1}{n_{p_1+1}^2} \left[ \frac{1}{n_1^2} \right] \frac{1}{\lambda_{\max}(\Sigma_0^{-1} \Sigma_1)}^2
\]

by definition of \( \lambda \) and by Lemma B.9,

\[
\leq \frac{\lambda_{\max}(\Sigma_0^{-1} \Sigma_1)}{n_1^2} \frac{1}{\lambda_{\max}(\Sigma_0^{-1} \Sigma_1)}^2
\]

again by definition of characteristic root,

\[
\leq \frac{B}{n_1} \frac{1}{\sigma_0^2}
\]

by Propositions A.3.2 and A.3.10. The last expression converges to zero as \( n \to \infty \) because \( n_1^2 = \nu_1 \) and \( \nu_1 \) converges to infinity by Assumptions 4.2.1 and 4.2.3.

It remains to show that \( \text{Var}_0 \left\{ \frac{\partial^2 \lambda(\chi, \xi)}{\partial \tau_i \partial \tau_j} \bigg| \chi = \chi_{0n}, \xi = \xi_{0n} \right\} \to 0 \) as \( n \to \infty \), \( i, j = 0, 1, \ldots, p_1 \). But

\[
\text{Var}_0 \left\{ \frac{\partial^2 \lambda(\chi, \xi)}{\partial \tau_i \partial \tau_j} \bigg| \chi = \chi_{0n}, \xi = \xi_{0n} \right\} = \text{Var}_0 \left\{ \frac{1}{2n_1 n_j} \left[ \text{tr} \Sigma_0^{-1} \Sigma_1^{-1} \Sigma_0^{-1} \Sigma_1^{-1} \right] \left( \chi - \chi_{0n} \right) \right\}
\]

\[
-2(\chi - \chi_{0n})' \Sigma_0^{-1} \Sigma_1^{-1} \Sigma_0^{-1} \Sigma_1^{-1} (\chi - \chi_{0n})
\]
using the same substitutions as above,

\[
\frac{2}{n_1 n_j} \text{tr}(\Xi_0^{-1} \Sigma_0^{-1} \Sigma_0^{-1} \Xi_0) \leq \frac{2(\min[m_i, m_j])}{n_1 n_j} \lambda_{\max}^2 (\Xi_0^{-1} \Sigma_0^{-1} \Sigma_0^{-1} \Xi_0)
\]

by Lemma B.1,

because there are at most \(\min[m_i, m_j]\) nonzero characteristic roots of

\[
\Xi_0^{-1} \Sigma_0^{-1} \Xi_0^{2}
\]

by Lemma B.9,

\[
\leq \frac{2(\min[m_i, m_j])}{n_1 n_j} \lambda_{\max}^2 (\Sigma_0^{-1} \Sigma_0^{-1} \Sigma_0^{-1} \Sigma_0^{-1})
\]

by Lemma B.13,

\[
\leq \frac{2}{\sigma_{0i}^2 \sigma_{0j}^2} \cdot \frac{\min[m_i, m_j]}{n_1 n_j}
\]

by Lemma B.9 and Proposition A.3.2,

\[
\rightarrow 0
\]

by Proposition A.3.10.

Thus in all cases, \(\text{Var}_0 \left\{ \frac{\partial^2 \lambda}{\partial \psi_i \partial \psi_j} \bigg|_{\psi = \psi_0} \right\} \rightarrow 0\) as \(n \rightarrow \infty\), which was to be proved.
A.4.4. Verification of Conditions 3.3.1.vi and 3.3.1.iv--Convergence

in Probability to Zero of \( \frac{\partial^2 \lambda}{\partial \psi \partial \psi'} \bigg|_{\hat{\psi}_n, \hat{\psi} = \hat{\psi}_0} \) \quad \text{Uniformly for} \quad \hat{\psi}_n \in \mathcal{S}_b(\hat{\psi}_0)

\( \hat{\psi}_n \in \mathcal{S}_b(\hat{\psi}_0) \).

**Lemma A.4.4.** For \( \psi \) and \( \lambda(\psi, \psi') \) as defined in Section 4.5, and for any \( b > 0 \), if Conditions A.2.1 and A.3.1 are true, then

\[
\sup_{\hat{\psi}_n \in \mathcal{S}_b(\hat{\psi}_0)} \left| \frac{\partial^2 \lambda(\psi, \psi')}{\partial \psi_i \partial \psi_j} \right|_{\hat{\psi} = \hat{\psi}_0} \quad \text{as} \quad n \to \infty, \quad i, j = 1, 2, \ldots, p.
\]

**Proof.**

Let \( \hat{\psi}_n \) be any point in \( \mathcal{S}_b(\hat{\psi}_0) \). Then for the first set of derivatives,

\[
\frac{\partial^2 \lambda}{\partial \psi \partial \psi'} \bigg|_{\hat{\psi} = \hat{\psi}_0} = -\frac{1}{n\sigma_{1}^{2} + 1} \chi_{1}^{T} \chi^{-1}_{1}, \quad a = 0, 1.
\]

(Recall \( \sigma_{0}^{2} = \hat{\psi}_{0} \).)

It is clearly sufficient for these derivatives to show for any \( \xi_{1}, \xi_{2} \),

\[
p_{0} \chi_{1} \text{ such that } \xi_{1}^{T} \xi_{1} = \xi_{2}^{T} \xi_{2} = 1 \text{ that } \phi = \frac{1}{n\sigma_{1}^{2} + 1} \left| \frac{1}{n\sigma_{1}^{2} + 1} \xi_{1}^{T} \left( \chi_{1}^{T} - \sigma_{0}^{2} \right) \xi_{2} \right| \to 0
\]
independent of \( \psi_{1n} \). But \( T_{1}^{-1} - T_{0}^{-1} = T_{1}^{-1}(\Sigma_{0} - T_{1}T_{1}^{T})T_{1}^{-1} \) and \( \Sigma_{0}^{-1} = A^{-T}A^{-1} \) as usual; then

\[
\phi^2 = \left\| \frac{1}{n_{p_{1} + 1}} \varepsilon_{2}^{T}X'(T_{1}^{-1} - T_{0}^{-1})X_{2} \right\|^2
\]

\[
\leq (\varepsilon_{1}^{T}\varepsilon_{1})(\varepsilon_{2}^{T}\varepsilon_{2}) \left[ \frac{1}{n_{p_{1} + 1}} \lambda_{\text{max}}(X_{1}^{T}X_{1}) \right]^2 \lambda_{\text{max}}\left(\frac{A^{-T}A}{\Sigma_{0} - T_{1}T_{1}^{T}}\right)^2
\]

by Proposition A.3.4 and the fact that \( T_{1}^{-1} - T_{0}^{-1} \) is symmetric. But

\[
\lambda_{\text{max}}\left(\frac{A^{-T}A}{\Sigma_{0} - T_{1}T_{1}^{T}}\right)^2 \leq \max_{k=1,2,...,n} |\lambda_{k}(\frac{1}{n_{p_{1} + 1}}(\Sigma_{0} - T_{1}T_{1}^{T})| |^2
\]

by Lemmas B.11 and B.14,

\[
\leq \frac{h_{b}^2}{\min_{i=0,1,...,p_{1}} (n_{i} \sigma_{0i})^2}
\]

by Proposition A.3.2. \( \frac{1}{n_{p_{1} + 1}} \lambda_{\text{max}}(X_{1}^{T}X_{1}) \) is bounded by A.3.10 and

thus \( \phi \leq \frac{h_{b}^2}{\min_{i=0,1,...,p_{1}} (n_{i} \sigma_{0i})^2} \) (B is taken from Proposition A.3.10) and hence

\( \phi \to 0 \) independent of \( \psi_{1n} \). (Since none of the bounds or convergence rates depend on \( \psi_{1n} \).)
For the next set of derivatives,

\[
\frac{\partial^2 \lambda}{\partial \tau_1 \partial \delta} \bigg|_{\delta = \delta_{an}} = \frac{1}{n_1 n_{p1+1}} \left( X' \tilde{T}_0^{-1} \tilde{G}_1 \tilde{T}_0^{-1} \left( X - \frac{\delta}{n_{p1+1}} \right) \right),
\]

\[i = 0, 1, \ldots, p_1, a = 0, 1.\]

Therefore, it is sufficient to show for these derivatives that for all \(\delta \neq p_0 X_l\) such that \(\delta' \delta = 1\) that

\[
\left| \frac{1}{n_1 n_{p1+1}} \left( X' \tilde{T}_0^{-1} \tilde{G}_1 \tilde{T}_0^{-1} \left( X - \frac{\delta}{n_{p1+1}} \right) - \frac{\beta_1}{\zeta_{p1+1}} \right) \right| \to 0
\]

independent of \(\zeta_{lin}\). Lemma B.16 applies here and thus it is sufficient to show that each of the following two terms goes to zero independent of \(\zeta_{lin}\).

\[
\phi_1 = \frac{1}{n_1 n_{p1+1}} \left| X' \tilde{T}_0^{-1} \tilde{G}_1 \tilde{T}_0^{-1} \left( X - \frac{\delta}{n_{p1+1}} \right) \right|
\]

(Recall that \(\frac{\beta_0}{n_{p1+1}} = \zeta_0\).)

\[
\phi_2 = \frac{1}{n_1 n_{p1+1}} \left| X' \tilde{T}_0^{-1} \tilde{G}_1 \tilde{T}_0^{-1} \left( \frac{\delta - \beta_1}{\zeta_{p1+1}} \right) \right|
\]

Now

\[
\phi_2^2 \leq \frac{1}{n_1} (\delta' \delta) (\delta_0 - \beta_1)'(\delta_0 - \beta_1) \left[ \frac{1}{n_{p1+1}} \lambda_{max} \left( X' \tilde{T}_0^{-1} X \right) \right]^2
\]
\[ \lambda_{\max}(A^T T_1^{-1} G_1 T_1^{-1} A) \leq \frac{1}{n_1^2} \cdot 1 \cdot p_0 b^2 B^2 \cdot \frac{1}{\sigma_{0i}} \cdot \frac{1}{2} \cdot \lambda_{\max}(E_0 T_1^{-1}) \]

by Propositions A.3.1, A.3.2, A.3.3, A.3.10,

\[ \leq \frac{1}{n_1^2} p_0 b^2 B \frac{1}{\sigma_{0i}} \frac{1}{2} 16 \]

\[ \rightarrow 0, \ i=0,1,\ldots,p_1, \]

independent of \( \Psi_{ln} \).

Now a technique is developed which is used often in this and subsequent sections. \( T_{1}^{-1} = \Sigma_1^{-1} + T_{1}^{-1} - \Sigma_0^{-1} = \Sigma_1^{-1} + \bar{A}, \) where \( \bar{A} = T_{1}^{-1} - \Sigma_0^{-1} \).

As noted previously, \( \bar{A} \) is symmetric and \( \bar{A} = T_{1}^{-1} (\Sigma_0^{-1} - T_{1}^{-1} \Sigma_0^{-1}) \), thus

\[ T_{1}^{-1} G_1 T_{1}^{-1} - \Sigma_0^{-1} G_1 \Sigma_0^{-1} = \Sigma_0^{-1} G_1 A + \Sigma_0^{-1} G_1 A + \Sigma_0^{-1} A \cdot \phi_1 \] then breaks down into three terms, each of which fits into the following formula from Proposition A.3.4 with appropriate choice of \( E_{1} \) and \( E_{2} \):

\[ \left[ \frac{1}{n_1 n_{p_1}^2 + 1} \right]^2 \leq \frac{1}{n_1^2} \cdot \frac{1}{n_{p_1}^2} \cdot \lambda_{\max}(\bar{A}) \cdot \lambda_{\max}(A^T T_{1}^{-1} G_1 T_{1}^{-1} A) \]

\[ \cdot (\Psi_{0} - \Psi_{0})^T \bar{A}^{-1} E_{2} (\Psi_{0} - \Psi_{0}) \]
\[ \frac{1}{n_1} \sum_{i=1}^{n_1} \left( \mathbf{X}_i - \mathbf{X}_0 \right) \left( \mathbf{X}_i - \mathbf{X}_0 \right)^T \leq \frac{1}{n_1} \mathbf{B} \lambda_{\max} \left( \mathbf{A'} \mathbf{F}_1 \mathbf{A} \mathbf{A'} \mathbf{F}'_1 \mathbf{A} \mathbf{A'} \mathbf{F}'_1 \mathbf{A} \right) \left( \mathbf{X}_i - \mathbf{X}_0 \right) \left( \mathbf{X}_i - \mathbf{X}_0 \right)^T \leq \frac{1}{n_1} \mathbf{B} \lambda_{\max} \left( \mathbf{A'} \mathbf{F}_2 \mathbf{A} \mathbf{A'} \mathbf{F}'_2 \mathbf{A} \mathbf{A'} \mathbf{F}'_2 \mathbf{A} \right) \left( \mathbf{X}_i - \mathbf{X}_0 \right) \left( \mathbf{X}_i - \mathbf{X}_0 \right)^T \]

by Proposition A.3.10. Now fit in the \( \frac{1}{n_1} \) where appropriate with \( \mathbf{F}_1 \) and \( \mathbf{F}_2 \) to obtain the division shown in Table A.4.4.1.

### Table A.4.4.1

Division of \( \phi_1 \) into Terms with Appropriate \( \mathbf{F}_1 \) and \( \mathbf{F}_2 \)

<table>
<thead>
<tr>
<th>Term</th>
<th>( \mathbf{F}_1 )</th>
<th>( \mathbf{F}_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \frac{1}{n_1} \sum_{i=1}^{n_1} \mathbf{G}_i \mathbf{G}_i^T )</td>
<td>( \sum_{i=1}^{n_1} \mathbf{G}_i \mathbf{G}_i^T )</td>
</tr>
<tr>
<td>2</td>
<td>( \sum_{i=1}^{n_1} \mathbf{G}_i \mathbf{G}_i^T )</td>
<td>( \frac{1}{n_1} \sum_{i=1}^{n_1} \mathbf{G}_i \mathbf{G}_i^T )</td>
</tr>
<tr>
<td>3</td>
<td>( \sum_{i=1}^{n_1} \mathbf{G}_i \mathbf{G}_i^T )</td>
<td>( \sum_{i=1}^{n_1} \mathbf{G}_i \mathbf{G}_i^T )</td>
</tr>
</tbody>
</table>

Proposition A.3.3 now yields the following bounds for \( \lambda_{\max} \left( \mathbf{A'} \mathbf{F}_1 \mathbf{A} \mathbf{A'} \mathbf{F}'_1 \mathbf{A} \mathbf{A'} \mathbf{F}'_1 \mathbf{A} \right) \)

which are given in Table A.4.4.2.
Table A.4.4.2

Bounds for $\lambda_{\text{max}}(A^\prime F_1 A A^\prime F_2 A)$ with $F_2$ from Table A.4.4.1

<table>
<thead>
<tr>
<th>Term</th>
<th>Bound for $\lambda_{\text{max}}(A^\prime F_1 A A^\prime F_2 A)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\frac{4}{n_1^2 \sigma_{0i}^2}$</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{4b^2}{\min_{j=0,1,\ldots,P_1} (n_j \sigma_{0j})^2}$</td>
</tr>
<tr>
<td>3</td>
<td>$\frac{16b^2}{n_1^2 \sigma_{0i}^2 \min_{j=0,1,\ldots,P_1} (n_j \sigma_{0j})^2}$</td>
</tr>
</tbody>
</table>

All these bounds go to zero independent of $\hat{y}_{1n}$; thus it is sufficient to show that $(y - y_{0})^\prime F_2 A^{-1} F_2 (y - y_{0})$ is bounded independent of $\hat{y}_{1n}$ for the two possible choices of $F_2$. The remarks preceding Proposition A.3.5 show that it is sufficient to show that $w^\prime Q^\prime A^\prime A^{-1} Q w$ is bounded for $s=0,1,\ldots,c$. But propositions A.3.5 and A.3.6 apply here and together with Conditions A.2.1 and A.3.1 they yield the following bounds.
Table A.4.4.3

Bounds for $\frac{\mathbf{w}_a \mathbf{A}^T \mathbf{A}^{-T} \mathbf{F} \mathbf{A}^{-1} \mathbf{F} \mathbf{A}_Q \mathbf{w}_a}{\mathbf{F}_\mathbf{X}}$ for $\mathbf{F}_\mathbf{X}$ from Table A.4.4.1

$$
\begin{align*}
\mathbf{F}_\mathbf{X} & \quad \text{Bound for } \frac{\mathbf{w}_a \mathbf{A}^T \mathbf{A}^{-T} \mathbf{F} \mathbf{A}^{-1} \mathbf{F} \mathbf{A}_Q \mathbf{w}_a}{\mathbf{F}_\mathbf{X}} \\
& \begin{cases}
0 & i \in S_{s+1}^* \\
\frac{1}{n_i} \mathbf{P}_{T_i}^{-1} & i \notin S_{s+1}^* \\
\frac{1}{2} \cdot \frac{n_i}{c_{0i}} & i \notin S_{s+1}^* \\
\frac{1}{10} \cdot \frac{n_i}{c_{0i}} \tilde{m}_s & i \notin S_{s+1}^* \\
\end{cases}
\end{align*}
$$

But Proposition A.3.10 guarantees that the bounds in Table A.4.4.3 are themselves bounded and hence for each $i=0,1,\ldots,p_1$ the desired convergence to zero independent of $\hat{\mathbf{F}}_{\mathbf{X}_1}$ takes place.

It now remains to treat the terms $\frac{\partial^2 \lambda}{\partial \tau_i \partial \tau_j}$.}

$$
\left. \frac{\partial^2 \lambda}{\partial \tau_i \partial \tau_j} \right|_{\hat{\mathbf{F}}_{\mathbf{X}_1}} = \frac{1}{2n_i n_j} \left[ \left( \text{tr} \left( \frac{\partial^2 \mathbf{G}_{i} \mathbf{G}_{j}}{\partial \mathbf{X}_1} \right) - 2(\mathbf{X} - \mathbf{X}) \frac{\partial^2 \mathbf{G}_{a}}{\partial \mathbf{X}_1} \right) \right] (\mathbf{X} - \mathbf{X}) + \frac{\partial^2 \mathbf{G}_{a}}{\partial \mathbf{X}_1} (\mathbf{X} - \mathbf{X}) \frac{\partial^2 \mathbf{G}_{a}}{\partial \mathbf{X}_1} \right],
$$

$i,j=0,1,\ldots,p_1$, $a=0,1$. 


It is sufficient for these derivatives to show that

\[ \phi_0 = \frac{1}{2n_1 n_j} \left| \text{tr} \frac{T_j^{-1} \Sigma_j T_j^{-1} \Sigma_j}{\Sigma_1 \Sigma_1} - \text{tr} \frac{\Sigma_0^{-1} \Sigma_0^{-1} \Sigma_0^{-1} \Sigma_0^{-1}}{\Sigma_1 \Sigma_1} \right| \]

and

\[ \phi_1 = \frac{1}{n_1 n_j} \left| \left( \frac{X - \Sigma}{n} \right)' \Sigma^{-1} \Sigma^{-1} \Sigma^{-1} \Sigma^{-1} \Sigma^{-1} \left( \frac{X - \Sigma}{n} \right) \right| \]

\[-\left( \frac{X - \Sigma}{n} \right)' \Sigma^{-1} \Sigma^{-1} \Sigma^{-1} \Sigma^{-1} \Sigma^{-1} \left( \frac{X - \Sigma}{n} \right) \]

each converge to zero independent of \( \Sigma \) as long as Conditions A.2.1 and A.3.1 are true.

For the first term write \( T_j^{-1} \Sigma_1 = \Sigma_0^{-1} + \Delta \). Since \( \text{tr} \Sigma + \text{tr} \Delta = \text{tr}(\Sigma + \Delta) \) for any matrices \( \Sigma \) and \( \Delta \), write \( \phi_0 \) as three terms and bound each separately. Each is of the form \( \frac{1}{2n_1 n_j} \left| \text{tr} \Sigma_1 \Sigma_1 \Sigma_1 \Sigma_1 \right| \) and hence

Propositions A.3.8 and A.3.9 apply to give the following bounds.
Table A.4.4.4

Bounds Used to Demonstrate Convergence to Zero of $\phi_0$

| Term | $\Sigma_0^{-1}$ | $T_1^{-1}(\Sigma_0^{-1} T_1) T_2^{-1} \Sigma_0^{-1}$ | Bound for $\frac{1}{2n_i n_j} |\text{tr} \ E_i G_i E_j G_j|$
|------|-----------------|-----------------------------------------------|--------------------------------------------------------------------------------|
| 1    | $T_1^{-1}(\Sigma_0^{-1} T_1) T_2^{-1} \Sigma_0^{-1}$ | $\min(m_i, m_j)$ | $\frac{b}{n_i n_j} \frac{\min_{k=0,1,\ldots,P_1} (n_k \sigma_{0k}) \sigma_{0i} \sigma_{0j}}{\sigma_{0i} \sigma_{0j}}$
| 2    | $T_1^{-1}(\Sigma_0^{-1} T_1) T_2^{-1} \Sigma_0^{-1}$ | $\min(m_i, m_j)$ | $\frac{b}{n_i n_j} \frac{\min_{k=0,1,\ldots,P_1} (n_k \sigma_{0k}) \sigma_{0i} \sigma_{0j}}{\sigma_{0i} \sigma_{0j}}$
| 3    | $T_1^{-1}(\Sigma_0^{-1} T_1) T_2^{-1} \Sigma_0^{-1}$ | $\min(m_i, m_j)$ | $\frac{b^2}{n_i n_j} \frac{\min_{k=0,1,\ldots,P_1} (n_k \sigma_{0k})^2 \sigma_{0i} \sigma_{0j}}{\sigma_{0i} \sigma_{0j}}$

Again Proposition A.3.10 guarantees that all these bounds converge to zero independent of $X_{4n}$ and hence $\phi_0 \to 0$ independent of $X_{4n}$.

Handling $\phi_1$ is fairly messy. Lemma B.16 applies giving four terms:

$$\phi_{11} = \frac{1}{2n_i n_j} |(X-X_{20})' [T_1^{-1} \Sigma_0^{-1} T_1^{-1} \Sigma_0^{-1} (X-X_{20})]|,$$

$$\phi_{12} = \frac{1}{n_i n_j} \frac{(E_i - E_j)}{n_{i+1}} X' T_1^{-1} \Sigma_0^{-1} T_1^{-1} \Sigma_0^{-1} (X-X_{20}) |.$$
\[
\phi_{13} = \frac{1}{n_1 n_j} \left| (x'x_0')^T \tilde{g}^{-1}_{1} \tilde{g}^{-1}_{j} \tilde{g}^{-1}_{j} x \right| \left( \tilde{g}_0 \tilde{g}_j \right) \frac{1}{n_{p_1 + 1}}
\]

\[
\phi_{14} = \frac{1}{2n_1 n_j} \left| \frac{\left( \tilde{g}_0 \tilde{g}_j \right)}{n_{p_1 + 1}} \right| \left( x' \tilde{g}^{-1}_{1} \tilde{g}^{-1}_{j} \tilde{g}^{-1}_{j} \tilde{g}_j x \right) \frac{\left( \tilde{g}_0 \tilde{g}_j \right)}{n_{p_1 + 1}}
\]

\[
\phi_{14} \text{ is easy to dispose of.}
\]

\[
\phi_{14}^2 \leq \frac{1}{2n_1 n_j} \left[ \left( \tilde{g}_0 \tilde{g}_j \right)' \left( \tilde{g}_0 \tilde{g}_j \right) \right] \left[ \frac{1}{n_{p_1 + 1}} \lambda_{\text{max}}(x'x_0 x)^2 \right]
\]

\[
\cdot \lambda_{\text{max}}(A'T_1 \tilde{g}_{j}^{-1} \tilde{g}_j^{-1} A' \tilde{g}_j^{-1} T_1 \tilde{g}_j^{-1})
\]

by Proposition A.3.4,

\[
\leq \frac{1}{2n_1 n_j} p_0 B_{2,6}^2 \frac{1}{\sigma_0^2}
\]

by Propositions A.3.3, A.3.2, A.3.1, and A.3.10. This certainly converges to zero independent of \(y_{ln}^2\).

Now use Proposition A.3.4 again with \(F_1 = \tilde{T}_1 \tilde{g}_{1}^{-1} T_1^{-1}\) (which is symmetric) and \(F_2 = G_j T_1^{-1}\) to obtain

\[
\phi_{12}^2 \leq \frac{1}{n_1 n_j} \left( \tilde{g}_0 \tilde{g}_j \right)' \left( \tilde{g}_0 \tilde{g}_j \right) \left[ \frac{1}{n_{p_1 + 1}} \lambda_{\text{max}}(x'x_0 x)^2 \right] \lambda_{\text{max}}(A'T_1 \tilde{g}_j^{-1} T_1^{-1})^2
\]

\[
\cdot (x'x_0)' \tilde{T}_1 \tilde{g}_{j}^{-1} A' \tilde{g}_j^{-1} T_1^{-1} (x'x_0)
\]
\[ \leq \frac{1}{n_1^2} \Sigma_{\Sigma_0}^2 \cdot \left( \frac{\lambda}{\sigma_r} \right)^2 (y - \bar{x}_2^{(0)})' \left[ \frac{T^{-1}}{n_j} \right] \Sigma_j^{-1} \left[ \frac{G_j^{-1}}{n_j} \right] (y - \bar{x}_2^{(0)}). \]

Since the first part clearly goes to zero it is sufficient to bound

\[ (y - \bar{x}_2^{(0)})' \left[ \frac{T^{-1}}{n_j} \right] \Sigma_j^{-1} \left[ \frac{G_j^{-1}}{n_j} \right] (y - \bar{x}_2^{(0)}) \quad \text{for} \quad F_2 = \frac{1}{n_j} \Sigma_j^{-1}. \]

But since Conditions A.2.1 and A.3.1 are true, the remarks following Proposition A.3.7 show that all the necessary bounds were obtained while working on the terms for

\[ \frac{\partial^2 \lambda}{\partial \tau \partial \phi}; \]

these bounds are contained in Table A.4.4.3. This all guarantees that \( \phi_{12} \to 0 \) as \( n \to \infty \) independent of \( \bar{\Sigma}_0 \) as long as Conditions A.2.1 and A.3.1 are true. \( \phi_{13} \) obviously follows the same bounds as \( \phi_{12} \) and hence also converges to zero.

To handle \( \phi_{11} \) use the fact that \( T^{-1} = T^{-1} + A \) and hence

\[ \frac{T^{-1}}{n_1} \Sigma_j^{-1} \left[ \frac{T^{-1}}{n_j} \right] - \frac{T^{-1}}{n_1} \Sigma_j^{-1} \left[ \frac{T^{-1}}{n_j} \right] \]

breaks into seven terms of which two typical ones are \( \frac{T^{-1}}{n_1} \Sigma_j^{-1} \left[ \frac{T^{-1}}{n_j} \right] \) and \( \frac{T^{-1}}{n_1} \Sigma_j^{-1} \left[ \frac{T^{-1}}{n_j} \right] \). Each of the seven terms that \( \phi_{11} \) breaks into are then of the form \( \frac{1}{2} (y - \bar{x}_2^{(0)})' \left[ \frac{T^{-1}}{n_1} \right] \Sigma_j^{-1} \left[ \frac{T^{-1}}{n_j} \right] (y - \bar{x}_2^{(0)}) \) for the following values of \( F_2, F_1, F_0 \) (again the \( n_1 \) and \( n_j \) are inserted in the appropriate places).
Table A.4.4.5
Division of $\phi_{11}$ into Seven Terms with Appropriate Choices of $F_0$, $F_1$, and $F_2$

<table>
<thead>
<tr>
<th>Term</th>
<th>$F_0$</th>
<th>$F_1$</th>
<th>$F_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\frac{1}{n_1} \Sigma_0^{-1} G_1$</td>
<td>$T_1^{-1}(\Sigma_0^{-1} T_1) \Sigma_0^{-1}$</td>
<td>$\frac{1}{n_j} F_j \Sigma_0^{-1}$</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{1}{n_1} \Sigma_0^{-1} G_1$</td>
<td>$\frac{1}{n_j} \Sigma_0^{-1} G_j T_1^{-1}$</td>
<td>$(\Sigma_0^{-1} T_1) \Sigma_0^{-1}$</td>
</tr>
<tr>
<td>3</td>
<td>$\Sigma_0^{-1}(\Sigma_0^{-1} T_1)$</td>
<td>$\frac{1}{n_1} T_1^{-1} G_1 \Sigma_0^{-1}$</td>
<td>$\frac{1}{n_j} F_j \Sigma_0^{-1}$</td>
</tr>
<tr>
<td>4</td>
<td>$\Sigma_0^{-1}(\Sigma_0^{-1} T_1)$</td>
<td>$\frac{1}{n_1 n_j} T_1^{-1} G_2 \Sigma_0^{-1} G_j T_1^{-1}$</td>
<td>$(\Sigma_0^{-1} T_1) \Sigma_0^{-1}$</td>
</tr>
<tr>
<td>5</td>
<td>$\Sigma_0^{-1}(\Sigma_0^{-1} T_1)$</td>
<td>$\frac{1}{n_1} T_1^{-1} G_3 \Sigma_0^{-1}(\Sigma_0^{-1} T_1) \Sigma_0^{-1}$</td>
<td>$\frac{1}{n_j} F_j \Sigma_0^{-1}$</td>
</tr>
<tr>
<td>6</td>
<td>$\frac{1}{n_1} \Sigma_0^{-1} G_4$</td>
<td>$\frac{1}{n_j} T_1^{-1}(\Sigma_0^{-1} T_1) \Sigma_0^{-1} G_j T_1^{-1}$</td>
<td>$(\Sigma_0^{-1} T_1) \Sigma_0^{-1}$</td>
</tr>
<tr>
<td>7</td>
<td>$\Sigma_0^{-1}(\Sigma_0^{-1} T_1)$</td>
<td>$\frac{1}{n_1 n_j} T_1^{-1} G_4 T_1^{-1}(\Sigma_0^{-1} T_1) \Sigma_0^{-1} G_j T_1^{-1}$</td>
<td>$(\Sigma_0^{-1} T_1) \Sigma_0^{-1}$</td>
</tr>
</tbody>
</table>
By inspection of Table A.4.4.5, it is easily seen that all $F_0$ and $F_2$ are of such a form as to fit into Table A.4.4.3. Furthermore all $F_1$ are such that $\lambda_{\max}(A'F_1A) = 0$. (This is done by application of Lemma B.13 and Proposition A.3.2 and then using Table A.4.4.2.) These two facts together yield that each of the seven terms converges to zero and hence that $\phi_{ll} \to 0$ as $n \to \infty$ independent of $\psi_{ln}$.

This now covers all possible cases of $\frac{\delta^2 \lambda}{\delta \psi_i \delta \psi_j}$ and thus the lemma is proved. |||

A.4.5. Verification of Conditions 3.3.1.v and 3.3.1.i--Uniform

Continuity of $\frac{\delta^2 \lambda}{\delta \psi_i \delta \psi_j}$ for $\psi_n \in S_b(\psi_{On})$ in Probability

Lemma A.4.5. For $\psi$ and $\lambda(\psi, \psi)$ as defined in Section 4.5 and for any $b > 0$, if Conditions A.2.1 and A.3.1 are true, then

$\frac{\delta^2 \lambda(\psi, \psi)}{\delta \psi_i \delta \psi_j}$ is a uniformly continuous function of $\psi_n$ in $S_b(\psi_{On})$, $i,j=1,2,\ldots,p.$

Proof.

Let $\eta > 0$ be given and let $\psi_{ln} \in S_b(\psi_{On})$. It must be shown that there exists $\delta > 0$ (without loss of generality $\delta < \frac{b}{2}$) such that for $\psi_{2n} \in S_\delta(\psi_{ln})$
\[
\frac{\partial^2 \lambda(X, \tilde{Y})}{\partial \tilde{y}_i \partial \tilde{y}_j} \bigg|_{\tilde{y} = \tilde{y}_{2n}} - \frac{\partial^2 \lambda(X, \tilde{Y})}{\partial \tilde{y}_i \partial \tilde{y}_j} \bigg|_{\tilde{y} = \tilde{y}_{1n}} < \eta
\]

for all \(i, j = 1, 2, \ldots, p\). Clearly it is sufficient to bound the various parts into which these expressions may break up by a constant times a function of \(\delta\) which decreases as \(\delta\) decreases. As in previous lemmata

\[
\tilde{\psi}_n = (\beta'_a, \tau_a, \tau_{ap_1}, \ldots, \tau_{ap_1}), \quad T_a = \Sigma_{i=0}^{p_1} \frac{\tau_{a_i}}{n_i} G_i, \text{ for } a = 1, 2. \quad T_0 = \Sigma_0;
\]

\[
\frac{\beta_0}{n_{p_1+1}} = \alpha_0.
\]

The derivatives to be considered first are

\[
\frac{\partial^2 \lambda}{\partial \tilde{y}_a \partial \tilde{y}_b} \bigg|_{\tilde{y} = \tilde{y}_{an}} = \frac{1}{n_{p_1+1}} \tilde{\psi}'^T \tilde{\psi}^{-1} \tilde{\psi}, \quad a = 0, 1, 2.
\]

It is clearly sufficient for these derivatives to show that for any

\(\tilde{\gamma}_1, \tilde{\gamma}_2, \tilde{\beta}_0 \times 1\) such that \(\tilde{\gamma}_1 \tilde{\gamma}_1 = \tilde{\gamma}_2 \tilde{\gamma}_2 = 1\), that \(\phi = \frac{1}{2n_{p_1+1}} |\tilde{\gamma}' \tilde{\gamma} (T_{22}^{-1} - T_{11}) \tilde{\gamma} \tilde{\gamma}_2|\)

can be bounded properly. But

\[
\phi^2 \leq (\tilde{\gamma}_1 \tilde{\gamma}_1)(\tilde{\gamma}_2 \tilde{\gamma}_2) \left[ \frac{1}{n_{p_1+1}} \lambda_{\text{max}}(X' \tilde{\Sigma}_0^{-1} X) \right]^2 \lambda_{\text{max}} \left[ A' (T_{22}^{-1} - T_{11}) A \right]^2
\]

by Proposition A.3.4 and the fact that \(T_{22}^{-1} - T_{11}^{-1}\) is symmetric. But
\[ T_{2}^{-1} = T_{1}^{-1} + T_{2}^{-1} \quad \text{where } A_{21} = T_{2}^{-1}A_{21}T_{1}^{-1} \]

is symmetric; thus

\[ \lambda_{\max}(A'_{2}T_{2}^{-1} - T_{2}^{-1}A)^{2} \]

\[ = \lambda_{\max}(A'_{2}A_{21}A) \]

\[ = \lambda_{\max}(A_1'A_2A_{21}A_1'A_2) \]

\[ = \lambda_{\max}(A_1'A_2^{-1}A_1A_2A_1A_2^{-1}A_1A_2^{-1}A_1A_2^{-1}) \]

\[ \leq \lambda_{\max}(A_1'A_2^{-1}A_1A_2) \lambda_{\max}(A_1'A_2^{-1}A_1A_2) \max_{k=1,2,\ldots,n} \left| \lambda_k \left[ \Sigma^{-1} - (I_{21} - I_{22}) \right] \right|^2 \]

by Lemmas 8, 11, 8, 14,

\[ \leq 4 \cdot 16 \cdot \frac{\delta^2}{\min_{i=0,1,\ldots,p_1} (n_i \sigma_0)^2} \]

by Proposition A.3.2. Combining all the above, note that

\[ \left( \frac{1}{n_{p_1+1}} \lambda_{\max} \left( \Sigma^{-1} - \chi \right) \right)^2 \leq B^2 \]

by Proposition A.3.10 and hence

\[ \phi^2 \leq \frac{64B^2}{\min_{i=0,1,\ldots,p_1} (n_i \sigma_0)^2} \delta^2 . \]

This is just of the form desired because there certainly exists a
constant such that \( \frac{64B^2}{\min_{i=0,1,\ldots,p_1} (n_i \sigma_{01})^2} \) is less than that constant for all \( n \).

For the next set of derivatives,

\[
\frac{\partial^2 \lambda}{\partial \tau_1 \partial \tau_2} \bigg|_{\tau_1=\tau_2=\tau_{an}} = \frac{1}{n_1 n_{p_1+1}} \left( X_1 \right) \left( \frac{\beta_a}{n_{p_1+1}} \right), \quad i=0,1,\ldots,p_1, \quad a=1,2.
\]

Therefore it is sufficient to show for these derivatives that for all \( \xi p_0 x \) such that \( \xi^T \xi = 1 \) that

\[
\frac{1}{n_1 n_{p_1+1}} |\xi^T \left( x Z_1 x^{-1} \left( y - X \frac{\beta_a}{n_{p_1+1}} \right) \right)|
\]

is suitably bounded. Lemma B.15 applies here resulting in three terms to bound.

\[
\phi_1 = \frac{1}{n_1 n_{p_1+1}} \left| \xi^T \left( x Z_1 x^{-1} \left( y - X \frac{\beta_a}{n_{p_1+1}} \right) \right) \right|
\]

\[
\phi_2 = \frac{1}{n_1 n_{p_1+1}} \left| \xi^T \left( x Z_1 x^{-1} \left( y - X \frac{\beta_a}{n_{p_1+1}} \right) \right) \right|
\]

\[
\phi_3 = \frac{1}{n_1 n_{p_1+1}} \left| \xi^T \left( x Z_1 x^{-1} \left( y - X \frac{\beta_a}{n_{p_1+1}} \right) \right) \right|
\]

Consider \( \phi_3 \) first.
\[ \phi_3^2 \leq \frac{1}{n_1} \left( \tilde{\xi}' \xi \right)' \left( \tilde{\xi}_2 - \tilde{\xi}_1 \right) \left[ \frac{1}{n_2} \lambda_{\max} \left( \frac{X' \Sigma_0^{-1} X}{n_2} \right) \right]^2 \lambda_{\max} \left( \frac{A'^{-1} G T^{-1} A}{n_2} \right)^2 \]

by Proposition A.3.4,

\[ \leq \delta^2 \frac{1}{n_1} p_0 \cdot 3^2 \cdot \frac{6^4}{\sigma_{01}^2} \]

by Propositions A.3.1, A.3.2, A.3.10. This is again of the desired form.

For \( \phi_2 \) as above \( T_{22}^{-1} = T_{21}^{-1} + A_{21} \) and thus there are three terms in the difference of the form \( \tilde{\xi}' \chi \left( \frac{\tilde{\xi}_0 - \tilde{\xi}_1}{n_2} \right) \) which can be bounded by Propositions A.3.4, A.3.2, A.3.3, A.3.1 and A.3.10 as follows in Table A.4.5.1.
Table A.4.5.1

Division of $\phi_2$ into Three Terms with Appropriate $F_1$ and Bounds for Squares of Each Term

<table>
<thead>
<tr>
<th>Term</th>
<th>$F_1$</th>
<th>Eventual Bound for its Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$T_1^{-1}(T_1 \cdot T_2)T_2^{-1}G_4T_1^{-1}T_1^{-1}$</td>
<td>$\frac{256p_0b^2}{n_1^2 \sigma_{0i} \min_{j=0,1,...,p_1} (n_j \sigma_{0j})^2} \cdot \delta$</td>
</tr>
<tr>
<td>2</td>
<td>$T_1^{-1}G_4T_2^{-1}(T_1 \cdot T_2)T_1^{-1}$</td>
<td>$\frac{256p_0b^2}{n_1^2 \sigma_{0i} \min_{j=0,1,...,p_1} (n_j \sigma_{0j})^2} \cdot \delta^2$</td>
</tr>
<tr>
<td>3</td>
<td>$T_1^{-1}(T_1 \cdot T_2)T_2^{-1}T_2^{-1}(T_1 \cdot T_2)T_1^{-1}$</td>
<td>$\frac{4096p_0b^2}{n_1^2 \sigma_{0i} \min_{j=0,1,...,p_1} (n_j \sigma_{0j})^4} \cdot \delta^4$</td>
</tr>
</tbody>
</table>

All of these bounds have the desired form.

It remains to dispose of $\phi_1$. Again there will be three terms of the form $\sum' \chi \bar{F}_1 \bar{F}_2 (\chi \cdot \bar{\chi} \cdot \bar{\alpha})$ which can be bounded by Proposition A.3.1.
Table A.4.5.2

Division of $\phi_1$ into Three Terms with Appropriate

$F_1$ and $F_2$

<table>
<thead>
<tr>
<th>Term</th>
<th>$F_1$</th>
<th>$F_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\frac{1}{n_1} T^{-1}_1 (T_1 - T_2) T^{-1}_2$</td>
<td>$\frac{1}{n_1} G^{-1}_4 T^{-1}_2$</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{1}{n_1} T^{-1}_1 e_1 e_2$</td>
<td>$(T_1 - T_2) T^{-1}_1$</td>
</tr>
<tr>
<td>3</td>
<td>$\frac{1}{n_1} T^{-1}_1 (T_1 - T_2) T^{-1}_2 (T_2 - T_2) T^{-1}_1$</td>
<td>$(T_1 - T_2) T^{-1}_1$</td>
</tr>
</tbody>
</table>

Again it is apparent that all these terms will yield proper bounds even after the further division into $c^2$ terms using the $Q_s$ matrices and $\omega_s$ vectors. For illustration consider Term 2 above.

$$\lambda_{\max}(\tilde{A}^T \tilde{Q}''_{\omega} \tilde{A}^T \tilde{Q}''_{\omega} \tilde{A}) \leq \frac{1}{n_1} \frac{64}{\sigma^2_{01}}$$

by Propositions A.3.3 and A.3.2. Further, as in the comments following Proposition A.3.7, $F_2 = (T_1 - T_2) T^{-1}_1 = (T_1 - T_2) (T^{-1}_2 + \Delta), (\Delta = T^{-1}_1 - T^{-1}_2)$.

Thus $\omega'_s Q'_s \tilde{A}^T E_2 \tilde{A}^{-1} T_2 \tilde{A} F_2 \tilde{A} Q_s \omega_s$ breaks into four terms and only two need be bounded. The first term is

$\omega'_s Q'_s \tilde{A}^{-1} (T_1 - T_2) \tilde{A}^{-1} (T_1 - T_2) Q_0^{-1} \tilde{A} Q_s \omega_s$; this is covered in
Proposition A.3.7 and has a bound involving $\delta^2$ and the proper denominator. The second term is

$$\nu_s' S_s' A' \Sigma_0^{-1}(\Sigma_0 - T_1)T_1^{-1}(T_1 - T_2)A^{-t} A^{-1}(T_1 - T_2)T_2^{-1}(\Sigma_0 - T_1)\Sigma_0^{-1} A S S_s' \nu_s$$

$$\leq \nu_s' S_s' A' \Sigma_0^{-1}(\Sigma_0 - T_1)A^{-t} A^{-1}(\Sigma_0 - T_1)\Sigma_0^{-1} A S S_s' \nu_s$$

$$\cdot \lambda_{\max}(A' T_1^{-1}(T_1 - T_2)A^{-t} A^{-1}(T_1 - T_2)T_2^{-1} A)$$

by definition of characteristic root. But the first term is that of Proposition A.3.6 and is properly bounded and the second is less than or equal to $\frac{4}{\min(n_j \sigma_{0j}^j)} \cdot \delta^2$. All the other cases also give proper bounds. (All terms are bounded by a correct function of $\delta$ and a suitable constant.) This settles the case for $\phi_1$ and hence for these derivatives of the form $\frac{\partial^2 \lambda}{\partial \tau_i \partial \phi_j}$.

Now consider the derivatives of the form

$$\frac{\partial^2 \lambda}{\partial \tau_i \partial \phi_j} \bigg|_{\phi = \phi_{an}} = \frac{1}{2n_i n_j} \left[ \left( \text{tr}_{T_a G_1 G_2 \Sigma_j^{-2}} \left( \Sigma_j^{-1} - \frac{\Sigma_j^{-1} \frac{\Sigma_j^{-1}}{T_a G_1 G_2 \Sigma_j^{-2}} - 2 \left( \Sigma_j^{-1} - \frac{\Sigma_j^{-1}}{T_a G_1 G_2 \Sigma_j^{-2}} \right) \left( \Sigma_j^{-1} - \frac{\Sigma_j^{-1}}{T_a G_1 G_2 \Sigma_j^{-2}} \right) \left( \Sigma_j^{-1} - \frac{\Sigma_j^{-1}}{T_a G_1 G_2 \Sigma_j^{-2}} \right) \left( \Sigma_j^{-1} - \frac{\Sigma_j^{-1}}{T_a G_1 G_2 \Sigma_j^{-2}} \right) \right) \right] \right],$$

$i, j = 0, 1, \ldots, p_1$, $a = 1, 2$.

It is sufficient for these derivatives to show that for all $i, j = 0, 1, \ldots, p_1$ that
\[ \phi_0 = \frac{1}{2n_i n_j} \left| \text{tr} T_{21}^{-1} G_{4i} T_{22}^{-1} G_{4j} - \text{tr} \frac{T_{21}^{-1} G_{4i} T_{21}^{-1} G_{4j}}{p_{1+1}} \right| \]

and

\[ \phi_1 = \frac{1}{n_i n_j} \left| \left( \frac{\partial}{\partial n_i} \frac{\partial}{\partial n_j} \right) \left( T_{21}^{-1} G_{4i} T_{22}^{-1} G_{4j} T_{22}^{-1} \left( \frac{\partial}{\partial n_i} \frac{\partial}{\partial n_j} \right) \right) \right| \]

are both bounded by suitable functions of \( \delta \).

For \( \phi_0 \) there are three terms of the form \( \frac{1}{2n_i n_j} \left| \text{tr} E_{4i} G_{4i} E_{4j} G_{4j} \right| \) and hence Propositions A.3.8, A.3.9 and A.3.10 apply to give the following bounds in Table A.4.5.3.

**Table A.4.5.3**

| Term | \( \frac{1}{2n_i n_j} \left| \text{tr} E_{4i} G_{4i} E_{4j} G_{4j} \right| \) | Bound for \( \phi_0 \) |
|------|-----------------|-----------------|
| 1    | \( T_{21}^{-1} (T_{21}^{-1} T_{22}) T_{22}^{-1} \) | \( \frac{8\delta}{\min(n_k \sigma_{0k} \sigma_{0i} \sigma_{0j})} \) |
| 2    | \( T_{21}^{-1} T_{22}^{-1} \) | \( \frac{8\delta}{\min(n_k \sigma_{0k} \sigma_{0i} \sigma_{0j})} \) |
| 3    | \( T_{21}^{-1} (T_{21}^{-1} T_{22}) T_{22}^{-1} \) | \( \frac{32\delta^2}{\min(n_k \sigma_{0k} \sigma_{0i} \sigma_{0j})} \) |

All these bounds are of the proper form.
Now $\phi_1$ breaks into nine parts by Lemma B.16 as follows:

\[
\phi_{11} = \frac{1}{n_1 n_j} |(x-x_0)'(T_2^{-1}g_4^{-1}T_2^{-1}g_3^{-1}T_3^{-1}g_4^{-1}T_2^{-1}) (x-x_0)|
\]

\[
\phi_{12} = \frac{1}{n_1 n_j} \left| \frac{(x-x_0)'}{n_{p_1+1}} x'(T_2^{-1}g_4^{-1}T_2^{-1}g_3^{-1}T_3^{-1}g_4^{-1}T_2^{-1}) (x-x_0)|
\]

\[
\phi_{13} = \frac{1}{n_1 n_j} |(x-x_0)'(T_2^{-1}g_4^{-1}T_2^{-1}g_3^{-1}T_3^{-1}g_4^{-1}T_2^{-1}) x \frac{(x-x_0)}{n_{p_1+1}}|
\]

\[
\phi_{14} = \frac{1}{n_1 n_j} \left| \frac{(x-x_0)'}{n_{p_1+1}} x'(T_2^{-1}g_4^{-1}T_2^{-1}g_3^{-1}T_3^{-1}g_4^{-1}T_2^{-1}) x \frac{(x-x_0)}{n_{p_1+1}}|
\]

\[
\phi_{15} = \frac{1}{n_1 n_j} |(x-x_0)'T_2^{-1}g_4^{-1}T_2^{-1}g_3^{-1}x \frac{(x-x_0)}{n_{p_1+1}}|
\]

\[
\phi_{16} = \frac{1}{n_1 n_j} \left| \frac{(x-x_0)'}{n_{p_1+1}} x'(T_2^{-1}g_4^{-1}T_2^{-1}g_3^{-1}T_3^{-1}) (x-x_0)|
\]

\[
\phi_{17} = \frac{1}{n_1 n_j} \left| \frac{(x-x_0)'}{n_{p_1+1}} x'(T_2^{-1}g_4^{-1}T_2^{-1}g_3^{-1}T_3^{-1}) x \frac{(x-x_0)}{n_{p_1+1}}|
\]

\[
\phi_{18} = \frac{1}{n_1 n_j} \left| \frac{(x-x_0)'}{n_{p_1+1}} x'(T_2^{-1}g_4^{-1}T_2^{-1}g_3^{-1}T_3^{-1}) x \frac{(x-x_0)}{n_{p_1+1}}|
\]

\[
\phi_{19} = \frac{1}{n_1 n_j} \left| \frac{(x-x_0)'}{n_{p_1+1}} x'(T_2^{-1}g_4^{-1}T_2^{-1}g_3^{-1}T_3^{-1}) x \frac{(x-x_0)}{n_{p_1+1}}|
\]
All of these terms can be properly bounded using Proposition A.3.4. The first four must be divided into seven terms in the usual way using $T^{-1}_{2} = T^{-1}_{1} + A_{21}$. The last three are easily bounded as follows using Propositions A.3.4, A.3.2, and A.3.10.

$$
\phi_{17}^2 \leq \frac{p_0^2 b^2}{n_1 n_j} \cdot b^2 \cdot \frac{4096}{\sigma_{01}^2 \sigma_{0j}^2} \delta^2
$$

$$
\phi_{18}^2 \leq \frac{p_0^2 b^2}{n_1 n_j} \cdot b^2 \cdot \frac{4096}{\sigma_{01}^2 \sigma_{0j}^2} \delta^2
$$

$$
\phi_{19}^2 \leq \frac{p_0^2}{n_1 n_j} \cdot b^2 \cdot \frac{4096}{\sigma_{01}^2 \sigma_{0j}^2} \delta^4
$$

All these bounds are of the proper form.

$\phi_{15}$ and $\phi_{16}$ are equal except for changing $i$ and $j$ so the same bound applies to each. Since a $\delta^2$ term will come from the $(\beta_{1} - \beta_{2})$ part it must be verified only that boundedness for the remainder can be derived by Proposition A.3.4 with $F_{1} = \frac{1}{n_1} T_{1}^{-1} G_{1} T_{1}^{-1}$ and $F_{2} = \frac{1}{n_j} G_{j} T_{j}^{-1}$.

But the same argument used above for $F_{2} = \frac{1}{n_j} G_{j} T_{j}^{-1}$ can also be applied to this $F_{1}$ yielding a proper bound. Thus $\phi_{15}$ and $\phi_{16}$ can be properly bounded.

Now the decomposition of the other terms is summarized; the source of the all important $\delta$ terms is noted. It can be verified by inspection
that all bounds will be of proper form. \( \phi_{14} \) breaks into seven terms of the form 

\[
\frac{(\beta_0 - \beta_1)'}{n_{p_{1+1}}} X'P_1X \frac{(\beta_0 - \beta_1)}{n_{p_{1+1}}} \]

with \( P_1 = E_1G_2E_0G_0E_3 \). Such terms are bounded by Propositions A.3.4, A.3.3, A.3.2, and A.3.10.

\[
\text{Table A.4.5.4}
\]

Division of \( \phi_{14} \) into Seven Terms with Appropriate \( E_1, E_2 \) and \( E_3 \) and Source of \( \delta \) Term

<table>
<thead>
<tr>
<th>Term</th>
<th>( E_1 )</th>
<th>( E_2 )</th>
<th>( E_3 )</th>
<th>Source of ( \delta ) Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \frac{T_{\tilde{z}<em>1}}{T</em>{\tilde{z}<em>2}} ) ( \frac{T</em>{\tilde{z}<em>2}}{T</em>{\tilde{z}<em>1}} ) ( T</em>{\tilde{z}<em>1}^{-1} ) ( T</em>{\tilde{z}<em>2}^{-1} ) ( T</em>{\tilde{z}<em>1}^{-1} ) ( T</em>{\tilde{z}_2}^{-1} )</td>
<td>( E_1 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>( \frac{T_{\tilde{z}<em>1}}{T</em>{\tilde{z}<em>2}} ) ( \frac{T</em>{\tilde{z}<em>2}}{T</em>{\tilde{z}<em>1}} ) ( T</em>{\tilde{z}<em>1}^{-1} ) ( T</em>{\tilde{z}<em>2}^{-1} ) ( T</em>{\tilde{z}<em>1}^{-1} ) ( T</em>{\tilde{z}_2}^{-1} )</td>
<td>( E_2 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>( \frac{T_{\tilde{z}<em>1}}{T</em>{\tilde{z}<em>2}} ) ( \frac{T</em>{\tilde{z}<em>2}}{T</em>{\tilde{z}<em>1}} ) ( T</em>{\tilde{z}<em>1}^{-1} ) ( T</em>{\tilde{z}<em>2}^{-1} ) ( T</em>{\tilde{z}<em>1}^{-1} ) ( T</em>{\tilde{z}_2}^{-1} )</td>
<td>( E_3 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>( \frac{T_{\tilde{z}<em>1}}{T</em>{\tilde{z}<em>2}} ) ( \frac{T</em>{\tilde{z}<em>2}}{T</em>{\tilde{z}<em>1}} ) ( T</em>{\tilde{z}<em>1}^{-1} ) ( T</em>{\tilde{z}<em>2}^{-1} ) ( T</em>{\tilde{z}<em>1}^{-1} ) ( T</em>{\tilde{z}_2}^{-1} )</td>
<td>( E_2, E_3 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>( \frac{T_{\tilde{z}<em>1}}{T</em>{\tilde{z}<em>2}} ) ( \frac{T</em>{\tilde{z}<em>2}}{T</em>{\tilde{z}<em>1}} ) ( T</em>{\tilde{z}<em>1}^{-1} ) ( T</em>{\tilde{z}<em>2}^{-1} ) ( T</em>{\tilde{z}<em>1}^{-1} ) ( T</em>{\tilde{z}_2}^{-1} )</td>
<td>( E_1, E_3 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>( \frac{T_{\tilde{z}<em>1}}{T</em>{\tilde{z}<em>2}} ) ( \frac{T</em>{\tilde{z}<em>2}}{T</em>{\tilde{z}<em>1}} ) ( T</em>{\tilde{z}<em>1}^{-1} ) ( T</em>{\tilde{z}<em>2}^{-1} ) ( T</em>{\tilde{z}<em>1}^{-1} ) ( T</em>{\tilde{z}_2}^{-1} )</td>
<td>( E_2, E_1 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>( \frac{T_{\tilde{z}<em>1}}{T</em>{\tilde{z}<em>2}} ) ( \frac{T</em>{\tilde{z}<em>2}}{T</em>{\tilde{z}<em>1}} ) ( T</em>{\tilde{z}<em>1}^{-1} ) ( T</em>{\tilde{z}<em>2}^{-1} ) ( T</em>{\tilde{z}<em>1}^{-1} ) ( T</em>{\tilde{z}_2}^{-1} )</td>
<td>( E_1, E_2, E_3 )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( \phi_{12} \) and \( \phi_{13} \) are the transpose of each other upon interchange of \( i \) and \( j \). \( \phi_{12} \) breaks into seven terms of the form 

\[
\frac{(\beta_0 - \beta_1)'}{n_{p_{1+1}}} X'P_1X \frac{(\beta_0 - \beta_1)}{n_{p_{1+1}}} \]

which can be bounded by Propositions A.3.1-A.3.7 and A.3.10.
Table A.4.5.5
Division of $\phi_{11}$ into Seven Terms with Appropriate $F_1$ and $F_2$ and Source of $\delta$ Term

<table>
<thead>
<tr>
<th>Term</th>
<th>$F_1$</th>
<th>$F_2$</th>
<th>Source of $\delta$ Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\frac{1}{n_1} T_{11}^{-1} (T_{12} - T_{22}) T_{22}^{-1} g_{11} T_{21}^{-1}$</td>
<td>$\frac{1}{n_1} g_{11} T_{21}^{-1}$</td>
<td>$F_1$</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{1}{n_1} T_{12}^{-1} g_{12} T_{22}^{-1}$</td>
<td>$\frac{1}{n_1} g_{12} T_{22}^{-1}$</td>
<td>$F_1$</td>
</tr>
<tr>
<td>3</td>
<td>$\frac{1}{n_1 n_j} T_{11}^{-1} g_{1j} T_{22}^{-1}$</td>
<td>$(T_{11} - T_{22}) T_{21}^{-1}$</td>
<td>$F_2$</td>
</tr>
<tr>
<td>4</td>
<td>$\frac{1}{n_1 n_j} T_{11}^{-1} g_{1j} T_{21}^{-1} (T_{12} - T_{22}) T_{22}^{-1} g_{2j} T_{21}^{-1}$</td>
<td>$(T_{11} - T_{22}) T_{21}^{-1}$</td>
<td>$F_1, F_2$</td>
</tr>
<tr>
<td>5</td>
<td>$\frac{1}{n_1 n_j} T_{11}^{-1} (T_{12} - T_{22}) T_{22}^{-1} g_{12} T_{22}^{-1} g_{2j} T_{22}^{-1}$</td>
<td>$(T_{11} - T_{22}) T_{21}^{-1}$</td>
<td>$F_1, F_2$</td>
</tr>
<tr>
<td>6</td>
<td>$\frac{1}{n_1} T_{11}^{-1} (T_{12} - T_{22}) T_{22}^{-1} g_{12} T_{21}^{-1} (T_{21} - T_{22}) T_{22}^{-1}$</td>
<td>$\frac{1}{n_1} g_{12} T_{21}^{-1}$</td>
<td>$F_1$</td>
</tr>
<tr>
<td>7</td>
<td>$\frac{1}{n_1 n_j} T_{11}^{-1} (T_{12} - T_{22}) T_{22}^{-1} g_{1j} T_{22}^{-1} (T_{11} - T_{22}) T_{22}^{-1} g_{2j} T_{22}^{-1}$</td>
<td>$(T_{11} - T_{22}) T_{21}^{-1}$</td>
<td>$F_1, F_2$</td>
</tr>
</tbody>
</table>

$\phi_{11}$ breaks into seven terms of the form $(x - x_0) F_2 F_1 F_2 (x - x_0)$

which can be bounded by Propositions A.3.1-A.3.7 and A.3.10.
Table A.4.5.6
Division of $\varphi_{11}$ into Seven Terms with Appropriate $F_0$, $F_1$, and $F_2$ and Source of $\delta$ Term

<table>
<thead>
<tr>
<th>Term</th>
<th>$F_0$</th>
<th>$F_1$</th>
<th>$F_2$</th>
<th>Source of $\delta$ Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$T_1^{-1} (T_{x1} - T_{x2})$</td>
<td>$\frac{1}{n_1} T_2^{-1} G_{x1} T_2^{-1}$</td>
<td>$\frac{1}{n_j} G_{xj} T_2^{-1}$</td>
<td>$F_0$</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{1}{n_1} T_1^{-1} G_{x1}$</td>
<td>$T_2^{-1} (T_{x1} - T_{x2}) T_2^{-1}$</td>
<td>$\frac{1}{n_j} G_{xj} T_2^{-1}$</td>
<td>$F_1$</td>
</tr>
<tr>
<td>3</td>
<td>$\frac{1}{n_1} T_1^{-1} G_{x1}$</td>
<td>$\frac{1}{n_j} T_1^{-1} G_{xj} T_2^{-1}$</td>
<td>$(T_1 - T_2) T_2^{-1}$</td>
<td>$F_2$</td>
</tr>
<tr>
<td>4</td>
<td>$\frac{1}{n_1} T_1^{-1} G_{x1}$</td>
<td>$\frac{1}{n_j} T_1^{-1} G_{xj} T_2^{-1}$</td>
<td>$(T_1 - T_2) T_2^{-1}$</td>
<td>$F_1 F_2$</td>
</tr>
<tr>
<td>5</td>
<td>$T_1^{-1} (T_{x1} - T_{x2})$</td>
<td>$\frac{1}{n_1 n_j} T_2^{-1} G_{x1} T_2^{-1} G_{xj} T_2^{-1}$</td>
<td>$(T_1 - T_2) T_2^{-1}$</td>
<td>$F_0 F_2$</td>
</tr>
<tr>
<td>6</td>
<td>$T_1^{-1} (T_{x1} - T_{x2})$</td>
<td>$\frac{1}{n_1} T_1^{-1} G_{x1} T_2^{-1} G_{xj} T_2^{-1}$</td>
<td>$\frac{1}{n_j} G_{xj} T_2^{-1}$</td>
<td>$F_0 F_1$</td>
</tr>
<tr>
<td>7</td>
<td>$T_1^{-1} (T_{x1} - T_{x2})$</td>
<td>$\frac{1}{n_1 n_j} T_2^{-1} G_{x1} T_2^{-1} (T_1 - T_2) T_2^{-1} G_{xj} T_2^{-1}$</td>
<td>$(T_1 - T_2) T_2^{-1}$</td>
<td>$F_0 F_1 F_2$</td>
</tr>
</tbody>
</table>

As noted before every term above yields a bound of the proper form after all decompositions (including the $\varphi_{22}$ decomposition) have been made. Each term also includes a $\delta$ term. Thus all terms have proper bounds and the lemma is true. |||
APPENDIX B

ALGEBRAIC LEMMAE USED IN PREVIOUS CHAPTERS

The following lemmae are used in previous chapters. No claim of
originality is made for any of them. They are collected here so that
the reader may easily refer to them. Only a few proofs are given.
The other lemmae can be proved by algebraic manipulation or applica-
tion of simple well known results. References are given where
appropriate.

LEMMA B.1. If \( \chi \sim \mathcal{N}_n(\mu, \Sigma) \) with \( \Sigma \) positive definite and \( B \) is an nxn
symmetric constant matrix and \( b \) an nx1 constant vector then

\[
\delta(\chi - b)'B(\chi - b) = \text{tr } B\Sigma + (\mu - b)'B(\mu - b)
\]

and

\[
\text{Var}(\chi - b)'B(\chi - b) = 2\text{tr}(B\Sigma)^2 + 4(\mu - b)'B\Sigma B(\mu - b)
\]

PROOF.

This lemma is easily proved by algebraic manipulation of the
moments up to fourth order of the multivariate normal distribution.
These moments can be found in Anderson (1958:39).

LEMMA B.2. If \( z \sim \mathcal{N}_n(0, I) \) and \( b \) is an nx1 constant vector and \( A \) is an
nxn diagonal constant matrix, then

\[
\delta\left\{ e^{i(\tilde{z}'z + \tilde{z}'Az)} \right\} = |I - 2iA|^{-\frac{1}{2}} e^{-\frac{1}{2}b'(I - 2iA)^{-1}b}.
\]

234
PROOF.

The proof of this lemma is merely an extension of proofs of the characteristic function of a multivariate normal random variable. It is done by completing the square. A reference is Plackett (1960:16-17).

**Lemma B.3.** If \( x_1, x_2, \ldots, x_p \) are all nonnegative and at least one \( x_i \) is positive and \( b_1, b_2, \ldots, b_p \) are all positive and \( a_1, a_2, \ldots, a_p \) have any sign,

\[
\min_{i=1,2,\ldots,p} \frac{a_i}{b_i} \leq \frac{\sum_{i=1}^{p} x_i a_i}{\sum_{i=1}^{p} x_i b_i} \leq \max_{i=1,2,\ldots,p} \frac{a_i}{b_i}.
\]

**Proof.**

Let \( \rho_i = \frac{x_i}{b_i} \). Then \( \sum_{j=1}^{p} x_j b_j \) is positive because at least one \( x_j \) is positive and all the \( b_j \) are. \( \) Then \( 0 \leq \rho_i \leq 1 \) and \( \sum_{i=1}^{p} \rho_i = 1 \).

Let \( c_i = \frac{a_i}{b_i} \). Then \( \sum_{i=1}^{p} \rho_i c_i \) is a weighted arithmetic mean of the \( c_i \) and is thus between \( \min c_i = \min \frac{a_i}{b_i} \) and \( \max c_i = \max \frac{a_i}{b_i} \). But

\[
\sum_{i=1}^{p} \rho_i c_i = \sum_{i=1}^{p} \frac{x_i}{b_i} \cdot \frac{a_i}{b_i} = \frac{\sum_{i=1}^{p} x_i a_i}{\sum_{i=1}^{p} x_i b_i} \quad \text{and the lemma is proved.} \]
LEMMA B.4. If $A$ is nxn positive definite and $B$ is nxn symmetric and if 
$\lambda_1(C) \geq \lambda_2(C) \geq \ldots \geq \lambda_n(C)$ are the characteristic roots of any nxn matrix 
$C$, and if $\mathbf{g}_1, \ldots, \mathbf{g}_n$ are any nx1 vectors, then 

$$
\lambda_1(A^{-1}B) \leq \sup_{\mathbf{x} \neq 0, \mathbf{x}^T \mathbf{g}_j = 0} \frac{\mathbf{x}^T \mathbf{Bx}}{\mathbf{x}^T \mathbf{Ax}} \quad j=1,2,\ldots,i-1
$$

and

$$
\lambda_1(A^{-1}B) \geq \inf_{\mathbf{x} \neq 0, \mathbf{x}^T \mathbf{g}_j = 0} \frac{\mathbf{x}^T \mathbf{Bx}}{\mathbf{x}^T \mathbf{Ax}} \quad j=1,2,\ldots,n-1
$$

PROOF:

The proof of the first statement is given in Anderson and 
Des Gupta (1963). The second statement is proved analogously. |||

LEMMA B.5. If $A$ is nxn positive definite and $B$ is nxn symmetric then 

$$
\lambda_1(A^{-1}B) \geq \inf_{\mathbf{x} \in \mathcal{S}} \frac{\mathbf{x}^T \mathbf{Bx}}{\mathbf{x}^T \mathbf{Ax}}
$$

where $\mathcal{S}$ is a linear space of dimension $i$.

PROOF.

$\mathcal{S}$ has a basis of $i$ vectors. Let them be $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_i$. This basis 
can be extended to a basis for $\mathbb{R}^n$ by adding $n-i$ more vectors orthogonal
to $\beta_1, \beta_2, \ldots, \beta_i$. Let these vectors be $\alpha_1, \alpha_2, \ldots, \alpha_{n-i}$. Then $x \in \mathcal{M}$ if and only if $x' \alpha_j = 0$ for $j = 1, 2, \ldots, n-i$. Lemma B.4 then applies, proving this lemma. \[\|\]

**Lemma B.6.** For $A$ and $B$ as in Lemma B.4, $\lambda_i(A^{-1}B) = \frac{x_i'Bx_i}{x_i'Ax_i}$ where $x_i$ is the characteristic vector associated with $\lambda_i$.

**Lemma B.7.** Suppose $B_1 = \sum_{i=0}^{p} b_{1i} G_i$, $B_0 = \sum_{i=0}^{p} b_{0i} G_i$, all $G_i$ are positive semidefinite and at least one is positive definite and $b_{0i} > 0$, $i = 0, 1, \ldots, p$. Then

\[\lambda_j(B_0^{-1}B_1) = \frac{\sum_{i=0}^{p} b_{1i} x_i'G_i x_j}{\sum_{i=0}^{p} b_{0i} x_i'G_i x_j}\]

where $x_j$ is the characteristic vector associated with $\lambda_j$,

\[\lambda_1(B_0^{-1}B_1) \leq \max_{i = 0, 1, \ldots, p} \frac{b_{1i}}{b_{0i}}.\]

\[\text{iii) If } H \text{ is an } nxm \text{ matrix such that } x'H = 0 \text{ implies } x'G_i x = 0, \text{ for some } j = 0, 1, \ldots, p \text{ then } \]

\[\lambda_{m+1}(B_0^{-1}B_1) \leq \max_{i = 0, 1, \ldots, j} \frac{b_{1i}}{b_{0i}}.\]
PROOF.

Statement i) follows from Lemma B.6 and the definitions of $B_0$ and $B_1$. Statement ii) is a simple application of Lemma B.3 to part i). To prove Statement iii), observe that

$$\lambda_{m+1}(B_0^{-1}B_1) \leq \sup_{x \neq 0, x' \neq 0} \frac{x'B_1x}{x'G_1x}$$

by Lemma B.4,

$$\leq \sup_{x \neq 0} \frac{\sum_{i=0}^{j} b_{i1} x'G_{i}x}{\sum_{i=0}^{j} b_{0i} x'G_{i}x}$$

$$\leq \max_{i=0,1,\ldots,j} \frac{b_{i1}}{b_{0i}}$$

by Lemma B.3. Note that in both parts ii) and iii) the correspondences between Lemma B.3 and Lemma B.7 are as follows:

<table>
<thead>
<tr>
<th>Lemma B.3</th>
<th>Lemma B.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_i$</td>
<td>$x'G_{i}x$</td>
</tr>
<tr>
<td>$a_i$</td>
<td>$b_{i1}$</td>
</tr>
<tr>
<td>$b_i$</td>
<td>$b_{0i}$</td>
</tr>
</tbody>
</table>
LEMMA B.8. Let $\mathbf{B}$ be an $n \times n$ symmetric matrix and $\mathbf{C}$ an $n \times m$ matrix, then

$$\max_{j=1,2,\ldots,m} |\lambda_j(C'C_B)| \leq \lambda_{\max}(C'C) \max_{k=1,2,\ldots,n} |\lambda_k(\mathbf{B})|.$$ 

PROOF.

$$\lambda_j(C'C_B) = \frac{x'_jC'Bx_j}{x'_jx_j}$$

$$= \frac{x'_jC'Bx_j}{x'_jC'x_j} \cdot \frac{x'_jC'x_j}{x'_jx_j},$$

where $x_j$ is the characteristic vector associated with $\lambda_j$. (If $Cx_j = 0$ that root is not of interest in any case since only the nonzero characteristic roots of $C'C_B$ are of concern. Therefore, since $x'_jC'Cx_j > 0$

$$|\lambda_j(C'C_B)| = \left| \frac{x'_jC'Bx_j}{x'_jC'x_j} \right| \frac{x'_jC'x_j}{x'_jx_j}$$

$$\leq \sup_{x \neq 0} \left| \frac{x'Bx}{x'x} \right| \lambda_{\max}(C'C)$$

$$= \lambda_{\max}(C'C) \max_{k=1,2,\ldots,n} |\lambda_k(\mathbf{B})|.$$ 

LEMMA B.9. Let $\mathbf{A}$ and $\mathbf{B}$ be $n \times n$ matrices with $\mathbf{A}$ nonsingular. Then

$$\lambda_i(\mathbf{AB}) = \lambda_i(\mathbf{BA}).$$
LEMMA B.10. Let \( B \) be an \( n \times n \) matrix. Then \( \lambda_1(B) = \lambda_1(B^\prime) \).

LEMMA B.11. Let \( A \) be an \( m \times n \) matrix and \( B \) be an \( n \times m \) matrix. Then the nonzero characteristic roots of \( AB \) are equal to the nonzero characteristic roots of \( BA \).

LEMMA B.12. Let \( x \) and \( y \) be \( n \times 1 \) vectors and \( A \) an \( n \times n \) matrix. Then
\[
(x'Ax)^2 \leq (x'y)(y'Ax) \lambda_{\max}(A^2).
\]

LEMMA B.13. Let \( A \) be \( n \times n \) positive semidefinite and let \( B \) be \( n \times n \). Then
\[
\max_{k=1,2,\ldots,n} |\lambda_k(AB)| \leq \lambda_{\max}(A) \max_{k=1,2,\ldots,n} |\lambda_k(B)|.
\]

PROOF.

This lemma follows immediately from Lemmas B.8 and B.9 and the fact that there exists \( C \) such that \( A \geq C \).

LEMMA B.14. For \( A \) \( n \times n \),
\[
\max_{k=1,2,\ldots,n} |\lambda_k(A^j)| = \max_{k=1,2,\ldots,n} |\lambda_k(A)|^j,
\]
\[
j=1,2,3,\ldots.
\]

LEMMA B.15. If \( S_1 \) and \( S_2 \) are \( n \times n \) matrices and \( t_0, t_1, t_2 \) and \( y \) are \( n \times 1 \) vectors, then the following statements are true.

1) \( S_2(y-t_2) - S_1(y-t_1) = (S_2-S_1)(y-t_0) + (S_2-S_1)(t_0-t_1) + S_2(t_1-t_2) \).
ii) \((x_{-t_2})'s_2(x_{-t_2}) - (x_{-t_1})'s_1(x_{-t_1})\)

\[= (x_{-t_0})'(s_{-t_2}-s_{-t_1}) + (x_{-t_0}-t_0)'(s_{-t_2}-s_{-t_1}) \]

\[+ (x_{-t_0})'(s_{-t_2}-s_{-t_1})(t_0-t_1) + (x_{-t_0}-t_0)'(s_{-t_2}-s_{-t_1})(t_0-t_1) \]

\[+ (x_{-t_0})'s_2(t_0-t_2) + (t_0-t_2)'s_2(x_{-t_0}) \]

\[+ (t_0-t_2)'s_2(t_0-t_2) \cdot \]

**Lemma B.16.** If \(S_1\) and \(S_0\) are nxn matrices and \(t_0, t_1\) and \(x\) are nx1 vectors, then the following statements are true.

i) \(S_1(x-t_1) - S_0(x-t_0) = (S_{-t_1}-S_0)(x-t_0) + S_1(t_0-t_1) \cdot \)

ii) \((x-t_1)s_1(x-t_1) - (x-t_0)s_0(x-t_0)\)

\[= (x_{-t_0})'(s_{-t_1}-s_0)(x_{-t_0}) \]

\[+ (t_0-t_1)'s_1(x_{-t_0}) \]

\[+ (x_{-t_0})'s_1(t_0-t_1) \]

\[+ (t_0-t_1)'s_1(t_0-t_1) \cdot \]
APPENDIX C

A COMPUTER PROGRAM TO IMPLEMENT

THE ITERATIVE PROCEDURE

C.1. Algebra Used in the Computer Program

A computer program has been written to implement The Iterative Procedure, which was introduced in Chapter 5. This section contains a brief description of the algebra used to write the program. This algebraic manipulation was used instead of more straightforward methods of calculation in order to reduce core storage requirements in the computer. The straightforward methods require the storage of several n×n matrices in order to compute the quantities required for the solution of the iterative equations (Σ and perhaps each of the Σ̃). Since this requires a great deal of core for even moderate n, some improvement is needed. The algebraic manipulations given here reduce the maximum dimension of the matrices which must be stored to

\[ m = \sum_{i=1}^{P_1} m_i \].

This is usually much smaller than n. For example, in the two-way balanced layout described in Section 6.1, n=IJK and m=IJ+I+J; even for small I,J,K appreciable savings can result. For instance, if I,J,K is 2,3,3 n=18 but m=11 and if I,J,K is 3,6,4 n=72 but m=27. The basic result used in these manipulations is due to Woodbury (1950) and is stated as Proposition C.1.1.

PROPOSITION C.1.1. Let \[ \Sigma = \Sigma_0 + UDU' \], where \[ \Sigma \] and \[ U \] are n×n and \[ U \] is
nmx and \( D \) is mnx diagonal and nonsingular. Then

\[
\Sigma^{-1} = \frac{1}{\sigma_0} \left( I - \Sigma (\sigma_0 \Sigma^{-1} + \Sigma')^{-1} \Sigma' \right).
\]

PROOF: This proposition is a special case of a result of Woodbury (1950). 

The matrix which must be inverted for this computer program may be written in the above form, where \( U = [U_1: U_2: \ldots : U_p] \) and

\[
D = \begin{bmatrix}
D_1 & 0 & \cdots & 0 \\
0 & D_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & D_p \\
\end{bmatrix}
\]

where \( D_1 = \sigma_1 I_{m_1} \). Thus \( m = \sum_{i=1}^{p} m_i \). Let \( \Sigma = \Sigma' \Sigma \) be partitioned as

\[
F_{\Sigma} = \begin{bmatrix}
F_{11} & \cdots & F_{1p} \\
\vdots & \ddots & \vdots \\
F_{p1} & \cdots & F_{pp}
\end{bmatrix}
\]
where $F_{ij} = U_i^t U_j$. Let $E = (\sigma_i D_i^{-1} + f)$ be partitioned similarly. Let

$$X_0 = U' X = \begin{bmatrix}
X_{(0)}^1 \\
X_{(0)}^2 \\
\vdots \\
X_{(0)}^p_1
\end{bmatrix},$$

where $X_{(0)}^1 = U_1^t X$. Let

$$H = U' X = \begin{bmatrix}
H_1 \\
H_2 \\
\vdots \\
H_{p_1}
\end{bmatrix},$$

where $H_1 = U_1^t X$. Let

$$A_0 = X' X$$

and

$$B_0 = X' X.$$
Then $E$ and $\bar{E}$ are $mxm$, $X_0$ is $mx1$, $H$ is $mxp_0$, $A_0$ is $p_0xp_0$ and $b_0$ is $p_0x1$.

These quantities and $\chi'\chi$ are the only quantities needed to compute an iteration of The Iterative Procedure. Note that $X_0$, $F$, $H$, $A_0$, $b_0$ and $\chi'\chi$ may all be computed as the data are read in, thus eliminating any need to save the large matrix $U$. $\bar{E}$ can then be formed at the start of each iteration and the matrices $Q = \bar{E}^{-1}F$ and $P = \bar{E}^{-1}H$ can be formed by solving linear equations. Thus it is never necessary to actually invert $E$.

The quantities necessary to perform one iteration of The Iterative Procedure are the elements of the matrix $B(\gamma(1))$ and the vector $\gamma[\gamma(1), \gamma(\gamma(1))]$ defined in Section 5.4. The iteration required is then

$$\gamma(i+1) = B^{-1}(\gamma(1)) \gamma[\gamma(1), \gamma(\gamma(1))],$$

which can also be solved without inverting the matrix $B(\gamma(1))$. Thus it is never necessary to invert any matrix to use The Iterative Procedure.

To obtain the elements of $\gamma$, $\gamma$ must first be calculated from the equation

$$\chi'\bar{E}^{-1}\chi \gamma = \chi'\bar{E}^{-1}\chi,$$

or

$$A \gamma = b.$$

But

$$A = \chi'\bar{E}^{-1}\chi = \frac{1}{\gamma_0} \chi'(I - UE^{-1}U')\chi$$
\[ \hat{a} = \frac{1}{\sigma_0} (X'X - X'\Sigma \Sigma' X) \]

Also

\[ b = X' \Sigma^{-1} X \]

\[ = \frac{1}{\sigma_0} X'(I - \Sigma \Sigma' \Sigma^{-1} X) \]

\[ = \frac{1}{\sigma_0} (X'X - X'\Sigma \Sigma' X) \]

\[ = \frac{1}{\sigma_0} (b_0 - E' \Sigma^{-1} X_0) \]

\[ \hat{a} \] is obtained as \( \hat{a} = A^{-1} b \) and the elements of \( \hat{a} \) are calculated from the

\[ [\hat{a}]_i = (X - X \hat{a})' \Sigma^{-1} \Sigma^{-1} (X - X \hat{a}), \ i = 0, 1, \ldots, p_1. \]

For \( i = 1, 2, \ldots, p_1 \) this reduces to

\[ (X - X \hat{a})' \Sigma^{-1} \Sigma^{-1} (X - X \hat{a}) \]

\[ = \frac{1}{\sigma_0^2} (X - X \hat{a})' (I - \Sigma \Sigma' \Sigma^{-1} X) X' (I - \Sigma \Sigma' \Sigma^{-1} X) (X - X \hat{a}) \]
\[ \begin{align*}
\tilde{t}_i &= \frac{1}{\sigma_0^2} \left[ U_d'(I-U'E^{-1}U')(\chi-\hat{\chi}) \right] \left[ U_d'(I-U'E^{-1}U')(\chi-\hat{\chi}) \right]' \\
&= \frac{1}{\sigma_0^2} t_i' t_i,
\end{align*} \]

where

\[ \tilde{t}_i = U_d'(I-U'E^{-1}U')(\chi-\hat{\chi}) . \]

But if \( \tilde{w} \) is defined by

\[ \tilde{w} = U' (\chi - \hat{\chi}) \]

\[ = \tilde{w}_0 - \hat{w} \]

\[ = \begin{bmatrix} \tilde{w}_1 \\ \tilde{w}_2 \\ \vdots \\ \tilde{w}_{p_1} \end{bmatrix} , \]

partitioned the same as \( \tilde{w}_0 \), then

\[ \begin{align*}
\tilde{t}_i &= U_d'(I-U'E^{-1}U')(\chi-\hat{\chi}) \\
&= U_d'(\chi - \hat{\chi}) - U_d'U'E^{-1}U'(\chi - \hat{\chi})
\end{align*} \]
\[
\begin{align*}
&= \frac{p_1}{\sum_{j=1}^{\mathcal{Q}_1} q_{1j} x_j} \\
&= \mathcal{P}_1 - x_i
\end{align*}
\]

where \( \mathcal{Q} \) and \( \mathcal{P} \) are partitioned just like \( \mathcal{E}_2 \) and \( \mathcal{F}_2 \) respectively. For \( i = 0 \) it is necessary to calculate

\[
(\chi - \hat{x})' \mathcal{Q}_1^{-1} \mathcal{Q}_2^{-1} (\chi - \hat{x})
\]

\[
= \frac{1}{\sigma_0} \left( (\chi - \hat{x})' (I - \mathcal{U} \mathcal{E}^{-1} \mathcal{U}') (I - \mathcal{U} \mathcal{E}^{-1} \mathcal{U}') (\chi - \hat{x}) - (\chi - \hat{x})' \mathcal{U} \mathcal{E}^{-1} \mathcal{U}' (\chi - \hat{x}) \right)
\]

\[
= \frac{1}{\sigma_0} \left( (\chi - \hat{x})' (\chi - \hat{x}) - 2 (\chi - \hat{x})' \mathcal{U} \mathcal{E}^{-1} \mathcal{U}' (\chi - \hat{x}) \right)
\]

\[
+ (\chi - \hat{x})' \mathcal{U} \mathcal{E}^{-1} \mathcal{U}' \mathcal{U} \mathcal{E}^{-1} \mathcal{U}' (\chi - \hat{x})
\]

\[
= \frac{1}{\sigma_0} \left( \chi' \chi - 2 \chi' \hat{x} + \hat{x}' \hat{x} - 2 \chi' \mathcal{U} \mathcal{E}^{-1} \mathcal{U}' (\chi - \hat{x}) \right)
\]

\[
+ \left[ \mathcal{U} \mathcal{E}^{-1} \mathcal{U}' (\chi - \hat{x}) \right]' \mathcal{U} \mathcal{E}^{-1} \mathcal{U}' (\chi - \hat{x})
\]

\[
= \frac{1}{\sigma_0} \left( \chi' \chi - 2 \hat{x}_0' \hat{x} + \hat{x}_0 \hat{x}_0 - 2 \chi' \hat{x} + \chi' \mathcal{E} \right)
\]

where \( z = \mathcal{E}^{-1} \mathcal{F} \) and \( z \) is partitioned as \( x_i \).
Now $[B]_{ij}$ must be computed, $i,j=0,1,\ldots,p_1$. For $i=j=0$,

$$[B]_{00} = \text{tr} \sum^{-1}_n \sum^{-1}_e$$

$$= \frac{1}{\sigma_0^2} \text{tr} \left( I - \sum^{-1}_n U' \right) \left( I - \sum^{-1}_e U' \right)$$

$$= \frac{1}{\sigma_0^2} \left\{ \text{tr} I - 2 \text{tr} \sum^{-1}_n U' + \text{tr} \sum^{-1}_n U' \sum^{-1}_e U' \right\}$$

$$= \frac{1}{\sigma_0^2} \left\{ n - 2 \text{tr} \sum^{-1}_n U' + \text{tr} \sum^{-1}_n U' \sum^{-1}_e U' \right\}$$

$$= \frac{1}{\sigma_0^2} \left\{ n - 2 \text{tr} \sum^{-1}_n + \text{tr} \sum^{-2}_n \right\}.$$  

For $j=0$, $i=1,2,\ldots,p_1$

$$[B]_{i0} = \text{tr} \sum^{-1}_n \sum^{-1}_e g_i$$

$$= \frac{1}{\sigma_0^2} \text{tr} \left( I - \sum^{-1}_n U' \right) \left( I - \sum^{-1}_e U' \right) g_{i,1}$$

$$= \frac{1}{\sigma_0^2} \left\{ \text{tr} g_{i,1} - 2 \text{tr} \sum^{-1}_n U' g_{i,1} + \text{tr} \sum^{-1}_n U' \sum^{-1}_e U' g_{i,1} \right\}$$

$$= \frac{1}{\sigma_0^2} \left\{ n - 2 \text{tr} \sum^{-1}_n U' g_{i,1} + \text{tr} \sum^{-1}_n U' \sum^{-1}_e U' g_{i,1} \right\}.$$
\[
= \frac{1}{\sigma_0^2} \left\{ n - 2 \sum_{k=1}^{P_1} \text{tr} F_{ik} Q_{kl} + \sum_{k=1}^{P_1} \sum_{l=1}^{P_1} \text{tr} Q_{ik} F_{kl} Q_{li} \right\}.
\]

For \( i, j = 1, 2, \ldots, P_1 \)

\[
[B]_{ij} = \text{tr} \Sigma^{-1} g_{\lambda \lambda} \Sigma^{-1} g_{\lambda \lambda}.
\]

\[
= \frac{1}{\sigma_0^2} \text{tr} (I - UE^{-1} U') U_i U_i' (I - UE^{-1} U') U_j U_j'.
\]

\[
= \frac{1}{\sigma_0^2} \text{tr} U_i' (I - UE^{-1} U') U_i U_i' (I - UE^{-1} U') U_j
\]

\[
= \frac{1}{\sigma_0^2} \text{tr} U_i' U_j - U_i' U_i U_j U_j' U_i U_i - U_i U_j U_j - U_i U_j U_j U_i
\]

\[
= \frac{1}{\sigma_0^2} \text{tr} [F_{ij} - \sum_{k=1}^{P_1} F_{ik} Q_{kj}] [F_{ij} - \sum_{k=1}^{P_1} F_{ik} Q_{kj}].
\]

The above formulae demonstrate that indeed only the matrices \( F_i, H, X_0, A_0 \) and \( b_0 \) described above need be carried in the calculations. Thus it is never necessary to save the large matrices \( U \) and \( X \). This enables problems of a reasonable size to be run using this program, which would not be possible if matrices of dimension \( n \) were needed.
The above algorithms are used in a sequence of subroutines which compute the requirements for each iteration. There is one more algebraic facet of the program of which the user should be aware. In Section 5.8 slowly converging sequences which oscillated above and below the final value were mentioned. This program has a feature designed to eliminate this problem. Whenever two iterates $g_{i-1}$ and $g_i$ are sufficiently different, $g_{i+1}$ is formed by taking another iteration from $g_i$ and then averaging. If averaging is necessary $g^{(1)}$ is calculated from $g_i$ as

$$g^{(1)} = \frac{g_i - g_{i-1}}{2} \cdot [g_i, g_i],$$

and then $g_{i+1}$ is formed as

$$g_{i+1} = \frac{1}{2} (g_i + g^{(1)}).$$

This feature eliminates the oscillating, but care must be taken that it not introduce a false convergence of its own. This is done by insisting that the last iteration must be one not involving the averaging process. The averaging and the safeguard process have worked very well on sample problems.
C.2. Setup and Output of the Computer Program

The computer program is designed to deliver as much freedom and convenience to the user as possible. The user must supply only two control cards. The first card contains the number of observations, \( n \), the number of levels of fixed factors, \( p_0 \), the number of random factors, \( p_1 \), a number indicating how often the iterates are to be printed (see below) and two yes-no statements. The first is yes if user supplies initial guesses and no otherwise. The second is yes if the short cut notation for the \( U \) matrix is used and no otherwise. (See below for explanation of the short cut notation.) The second control card contains the number of levels at which each random factor appears, \( m_i, i=1,2,\ldots,p_1 \). (Thus there are \( p_1 \) numbers on the second control card.)

After the control cards come the data cards. The data is read in by rows or observations. There are two cards per row. The first card contains the appropriate row of the \( U \) matrix. This may be in the form of zeroes and ones (there will be \( m = \sum_{i=1}^{p_1} m_i \) numbers.) or in short cut notation. In short cut notation only \( p_1 \) numbers are used, each number stating in which column of \( U \) the one appears; this is a unique description because by definition there is exactly one 1 in each row. The second card contains the appropriate row of the \( \chi \) matrix and the observation on \( y \). There must be \( n \) pairs of cards, one pair for each observation. The last card for a problem states yes or no—yes if another problem follows and no otherwise.
The output of the program consists first of the intermediate iterations the user asked to be printed. If the user placed a K on the first control card in the appropriate position, then the results after each K iterations are printed. Then follow the final results and the number of iterations required. The estimated large sample covariance matrices for \( \hat{\Sigma} \) and \( \hat{\Sigma} \) are also printed; these matrices are estimated by \( [X'X^{-1}(\hat{\Sigma})X]\^{-1} \) and \( 2B^{-1}(\hat{\Sigma}) \) respectively.

This computer program can handle fairly large problems with relative ease. The size of the largest matrix which should be inverted (or set of equations to be solved) is certainly less than 100 and probably closer to 50. (The reasons are twofold—computational accuracy and time requirements.) If \( \Sigma \) were inverted directly the size of problem would be severely limited. However, using the indirect methods above much larger problems can be accommodated. Since very often \( n \) is approximately a multiple of the product of some of the \( m_i \), even when the sum of the \( m_i \) is restricted to be small, \( n \) could be large. For instance, even if the sum of the \( m_i \) is less than 60, the possible values of \( n \) could be well over 1000. The simplicity of the control cards and data input makes the program easy to use even for the statistician who is unfamiliar with computer programming. It seems to be an effective and efficient program, judged by its performance on sample problems.

Copies of the program deck are available from the author upon request along with more detailed documentation and sample problems.
APPENDIX D

COMMENTS ON THE PROOF OF CONSISTENCY OF THE MAXIMUM
LIKELIHOOD ESTIMATES GIVEN BY HARTLEY AND RAO

In Section 1.1, it was mentioned that H. O. Hartley and J. N. K. Rao (1967) gave a proof of the consistency of the maximum likelihood estimates in the mixed model of the analysis of variance. It was remarked that the theorem was true but that the assumptions used were not all stated correctly, that the assumptions were restrictive and that important details were omitted from the proof. In this section some brief comments will be presented to elucidate these remarks. The notation used will be the notation used in this paper rather than that used by Hartley and Rao in order to preserve continuity.

The assumption stated incorrectly occurs on page 102 of the paper and is labeled 7.iii. It requires that for each design in the sequence of designs the matrix \( \sim [X:U_1:...:U_p] \) have a basis \( \sim [X:U^*] \) where \( U^* \) contains at least one column from each \( U_1 \). An assumption of this sort is necessary to insure estimability of the parameters. That this assumption is not sufficient can be seen in the following counterexample. Let \( U_1 \) be any matrix such that the basis of \( [X:U_1] \) is \( [X:U^*] \) where \( U^* \) contains at least \( p_1 \) columns from \( U_1 \). Then let \( U_j = U_1, j=2,3,...,p_1 \). In such a model instead of each individual \( \sigma_1, \sigma_2,...,\sigma_{p_1} \) only \( \sigma_1 + \sigma_2 + ... + \sigma_{p_1} \) can be estimated, yet this model satisfies the given assumption. One possible way to correctly state the assumption is to
use two separate assumptions like Assumptions 1.3.4 and 1.3.5.

The assumption which is restrictive is 7.1 (p. 101) which requires that all the diagonal elements of the matrix $\sum_{i} U_i U_i$ (of which there are $m_i$) be less than or equal to some universal constant $R$ for all $n$, $i=1,2,\ldots,p_1$. It was claimed in this paper that this required that the number of observations at any level of any random factor to be bounded and that it forced all normalizing sequences to be the order of $n$. That this assumption would be restrictive was shown in Section 1.1 by reference to Section 6.1. That the claims are true can be seen by the following argument. As Hartley and Rao point out, a diagonal element of $\sum_{i} U_i U_i$ represents exactly the number of observations at the appropriate level of the $i^{th}$ factor and the sum of these elements must be $n$. (Assumption 1.3.6, which was also used by Hartley and Rao (p.95), can be restated to say that every observation is allocated to exactly one level of each factor and that each level is allocated at least one observation.) Thus there are $m_i$ such elements, each less than $R$ (which demonstrates the first claim) and greater than zero, and adding up to $n$; it follows that $m_i \geq n/R$, $i=1,2,\ldots,p_1$, which was to be shown.

In an attempt to show that the details of the proof of consistency omitted by Hartley and Rao may be important, a short outline of the method of proof will be attempted. The object is to apply the method of Wald (1949) and Wolfowitz (1949) to $[L(\chi, \theta)]^\frac{1}{n}$, where $L(\chi, \theta)$ is the likelihood function. One shows that for each $\eta > 0$ and $\epsilon > 0$ there is an $h \equiv h(\eta)$ and an $n_0 \equiv n_0(\eta, \epsilon)$ such that $h < 1$ and such that
\[
\sup_{\theta \in \Omega(\eta)} \mathbb{P}_0 \left\{ \frac{\sum_{i=1}^{n} (\log L(x_i, \theta) - \log L(x_i, \theta_0))}{\log n} > \epsilon \right\} < \epsilon,
\]

where \(\Omega(\eta) = \{\theta : \|\theta - \theta_0\| > \eta\}\) and \(\theta_0\) is the true parameter point. (That is, the likelihood function cannot be large unless \(\theta\) is near \(\theta_0\).) This is done by showing that
\[
\sup_{\theta \in \Omega(\eta)} \mathbb{P}_0 \left\{ \frac{1}{n} \log L(x, \theta) - \frac{1}{n} \log L(x, \theta_0) > \log h \right\} < \epsilon.
\]

The following are some of the things which must be shown in order to use the Wald-Wolfowitz argument.

1. \(\frac{1}{n} \log L(x, \theta) \leq \mathbb{E}_0 \left[ \frac{1}{n} \log L(x, \theta) \right]\) for all \(\theta\).
2. \(\mathbb{E}_0 \left[ \frac{1}{n} \log L(x, \theta) - \frac{1}{n} \log L(x, \theta_0) \right] < 0\) for all \(n\).
3. \(\lim_{n \to \infty} \mathbb{E}_0 \left[ \frac{1}{n} \log L(x, \theta) - \frac{1}{n} \log L(x, \theta_0) \right] < 0\).
4. Continuity and limit conditions on \(L(x, \theta)\).

Hartley and Rao prove (1) as Lemma 7.1 (p.102). They state that (2) and (4) follow from the assumptions; this is true, although a great deal of work is necessary to prove (4). They do not mention (3) at all. However, Condition (3) is very important. It is necessary that the expected values in (2) approach a strictly negative limit in order to insure that the probability of any particular difference of log likeli-
hoods being negative will converge to one. To see that (1) and (2) above are not enough, consider the sequence of random variables 
\[ Z_n \sim \mathcal{N}\left(-\frac{1}{n}, \frac{1}{n^2}\right). \] Then \[ Z_n \xrightarrow{D} \delta Z_n \] and \[ \delta Z_n < 0 \] for all \( n \) but 
\[ P[Z_n < 0] = 0.8413 \] for all \( n \). Thus a condition like Condition (3) must be proved; to prove this condition it is necessary to use arguments of the same type used in Section A.4.1 to show that the matrix \( J \) was positive definite. It was shown in Section 1.1 that arguments about the positive definiteness of \( J \) were closely related to problems of degenerating distributions. That such arguments must also be taken into account in this version of a consistency proof is entirely reasonable since this result is a stronger result than Theorem 4.4.1 and so this version should take into account any difficulties that arise there. Condition (3) can be proved using all the assumptions—including 7.1, but it is not true when different parameters converge at different rates. Thus this proof of consistency can be used under all of Hartley and Rao's assumptions but will not work in the more general cases because the limit in (3) degenerates to zero.

Interestingly enough, although the assumptions 7.1-7.iii are necessary to prove consistency, they are not necessary to prove (1). In fact the only necessities are that the elements of \( \chi \) be bounded and that \( \theta \) be an interior point of the parameter space. To prove (1) it suffices to prove, as Hartley and Rao did in Lemma 7.1 (p.102), that 
\[ \text{Var}\left[\frac{1}{n} \log L(\chi, \theta)\right] = 0\left(\frac{1}{n}\right). \] But
\[
\log L(\chi, \theta) = \frac{n}{2} \log 2\pi - \frac{1}{2} \log |\Sigma| - \frac{1}{2}(\chi - \chi_0)^T \Sigma^{-1}(\chi - \chi_0).
\]

Therefore,

\[
\text{Var}_0[\log L(\chi, \theta)] = 2\text{tr}(\Sigma^{-1} \Sigma_0)^2 + 4(\chi - \chi_0)^T \Sigma^{-1} \Sigma_0^{-1} \Sigma^{-1} \Sigma_0 (\chi - \chi_0),
\]

by Lemma B.1,

\[
\leq 2n\lambda_{\max}^2(\Sigma^{-1} \Sigma_0) + 4(\chi - \chi_0)^T (\chi - \chi_0) \lambda_{\max}(\Sigma^{-1} \Sigma_0)
\]

\[
\cdot \lambda_{\max}(\Sigma^{-1} \Sigma_0)^2
\]

by Proposition A.3,4,

\[
\leq 2 \max_{i=0,1,\ldots,p_1} \left( \frac{\sigma_{ii}}{\sigma_{0i}} \right)^2 [n + 2(\chi - \chi_0)^T (\chi - \chi_0) \lambda_{\max}(\Sigma^{-1} \Sigma_0)]
\]

by Lemma B.7. But

\[
\lambda_{\max}(\Sigma^{-1} \Sigma_0) \leq \lambda_{\max}(\Sigma^{-1} \Sigma_0) \lambda_{\max}(\Sigma^{-1} \Sigma_0)
\]

by Lemma B.8

\[
= \lambda_{\max}(\Sigma^{-1} \Sigma_0) \frac{1}{\lambda_{\min}(\Sigma_0)}
\]

\[
\leq \lambda_{\max}(\Sigma^{-1} \Sigma_0) \frac{1}{\sigma_{00}}
\]

because \(\lambda_{\min}(\Sigma_0) \geq \sigma_{00}\). But if the elements of \(\Sigma\) are bounded then \(\lambda_{\max}(\Sigma^{-1} \Sigma_0) \leq K n\) for some constant \(K\). Therefore
\[ \text{Var}_0 \left[ \frac{1}{n} \log L(\tilde{y}, \tilde{\theta}) \right] \leq \frac{1}{n} \left[ 2 \max_{i=0,1,\ldots,p} \left( \frac{\sigma_{0i}}{\sigma_i} \right)^2 + 4(\tilde{\alpha} - \alpha_0)(\tilde{\alpha} - \alpha_0) \cdot \frac{K}{\sigma_{00}} \right] \]

which was to be proved.
BIBLIOGRAPHY


with linear structure", Proceedings of the Second International
Symposium on Multivariate Analysis (P. R. Krishnaiah, ed.),

Anderson, T. W. (1970), "Estimation of covariance matrices which are
one linear combination or whose inverses are linear combinations
of given matrices", Essays in Probability and Statistics, University


structure and moving average processes of finite order", Technical
Report No. 6, Stanford University.


Anderson, T. W. and S. Das Gupta (1963), "Some inequalities on character-

New York.

University Press, Princeton.


