Proof of Result in Chapter 10

Proof of Result 10.1 We assume that Σ_{11} and Σ_{22} are nonsingular.[1] Introduce the symmetric square-root matrices $\Sigma_{11}^{1/2}$ and $\Sigma_{22}^{1/2}$ with $\Sigma_{11} = \Sigma_{11}^{1/2}\Sigma_{11}^{1/2}$ and $\Sigma_{11}^{-1} = \Sigma_{11}^{-1/2}\Sigma_{11}^{-1/2}$. [See (2-22).] Set $\mathbf{c} = \Sigma_{11}^{1/2}\mathbf{a}$ and $\mathbf{d} = \Sigma_{22}^{1/2}\mathbf{b}$, so $\mathbf{a} = \Sigma_{11}^{-1/2}\mathbf{c}$ and $\mathbf{b} = \Sigma_{22}^{-1/2}\mathbf{d}$. Then

$$\operatorname{Corr}(\mathbf{a}'\mathbf{X}^{(1)}, \mathbf{b}'\mathbf{X}^{(2)}) = \frac{\mathbf{a}'\Sigma_{12}\mathbf{b}}{\sqrt{\mathbf{a}'\Sigma_{11}\mathbf{a}}\sqrt{\mathbf{b}'\Sigma_{22}\mathbf{b}}} = \frac{\mathbf{c}'\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2}\mathbf{d}}{\sqrt{\mathbf{c}'\mathbf{c}}\sqrt{\mathbf{d}'\mathbf{d}}} \tag{1}$$

By the Cauchy–Schwarz inequality (2-48),

$$\mathbf{c}'\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2}\mathbf{d} \le (\mathbf{c}'\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}\mathbf{c})^{1/2}(\mathbf{d}'\mathbf{d})^{1/2} \tag{2}$$

Since $\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}$ is a $p \times p$ symmetric matrix, the maximization result (2-51) yields

$$\mathbf{c}'\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}\mathbf{c} \le \lambda_1\mathbf{c}'\mathbf{c} \tag{3}$$

where λ_1 is the largest eigenvalue of $\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}$. Equality occurs in (3) for $\mathbf{c} = \mathbf{e}_1$, a normalized eigenvalue associated with λ_1. Equality also holds in (2) if $\mathbf{d}$ is proportional to $\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1/2}\mathbf{e}_1$. Thus,

$$\max_{\mathbf{a},\mathbf{b}} \operatorname{Corr}(\mathbf{a}'\mathbf{X}^{(1)}, \mathbf{b}'\mathbf{X}^{(2)}) = \sqrt{\lambda_1} \tag{4}$$

with equality occurring for $\mathbf{a} = \Sigma_{11}^{-1/2}\mathbf{c} = \Sigma_{11}^{-1/2}\mathbf{e}_1$ and with $\mathbf{b}$ proportional to $\Sigma_{22}^{-1/2}\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1/2}\mathbf{e}_1$, where the sign is selected to give positive correlation. We take $\mathbf{b} = \Sigma_{22}^{-1/2}\mathbf{f}_1$. This last correspondence follows by multiplying both sides of

$$(\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2})\mathbf{e}_1 = \lambda_1\mathbf{e}_1$$

by $\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1/2}$, yielding

$$\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1/2}(\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1/2}\mathbf{e}_1) = \lambda_1(\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1/2}\mathbf{e}_1) \tag{5}$$

Thus, if $(\lambda_1, \mathbf{e}_1)$ is an eigenvalue–eigenvector pair for $\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}$, then $(\lambda_1, \mathbf{f}_1)$—with $\mathbf{f}_1$ the normalized form of $\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1/2}\mathbf{e}_1$—is an eigenvalue–eigenvector pair for $\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1/2}$. The sign for $\mathbf{f}_1$ is chosen to give a positive correlation. We have demonstrated that $U_1 = \mathbf{e}_1'\Sigma_{11}^{-1/2}\mathbf{X}^{(1)}$ and $V_1 = \mathbf{f}_1'\Sigma_{22}^{-1/2}\mathbf{X}^{(2)}$ are the first pair of canonical variables and that their correlation is $\rho_1^* = \sqrt{\lambda_1}$. Also, $\operatorname{Var}(U_1) = \mathbf{e}_1'\Sigma_{11}^{-1/2}\Sigma_{11}\Sigma_{11}^{-1/2}\mathbf{e}_1 = \mathbf{e}_1'\mathbf{e}_1 = 1$, and similarly, $\operatorname{Var}(V_1) = 1$.

Continuing, we note that U_1 and an arbitrary linear combination $\mathbf{a}'\mathbf{X}^{(1)} = \mathbf{c}'\Sigma_{11}^{-1/2}\mathbf{X}^{(1)}$ are uncorrelated if

$$0 = \operatorname{Cov}(U_1, \mathbf{c}'\Sigma_{11}^{-1/2}\mathbf{X}^{(1)}) = \mathbf{e}_1'\Sigma_{11}^{-1/2}\Sigma_{11}\Sigma_{11}^{-1/2}\mathbf{c} = \mathbf{e}_1'\mathbf{c}, \quad \text{or} \quad \mathbf{c} \perp \mathbf{e}_1$$

At the kth stage, we require that $\mathbf{c} \perp \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{k-1}$. The maximization result (2–52) then yields

$$\mathbf{c}'\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}\mathbf{c} \le \lambda_k\mathbf{c}'\mathbf{c} \quad \text{for } \mathbf{c} \perp \mathbf{e}_1, \ldots, \mathbf{e}_{k-1}$$

[1] If Σ_{11} or Σ_{22} is singular, one or more variables may be deleted from the appropriate set, and the linear combinations $\mathbf{a}'\mathbf{X}^{(1)}$ and $\mathbf{b}'\mathbf{X}^{(2)}$ can be expressed in terms of the reduced set. If $p > \operatorname{rank}(\Sigma_{12}) = p_1$, then the nonzero canonical correlations are $\rho_1^*, \ldots, \rho_{p_1}^*$.

and by (1) ,

$$\operatorname{Corr}(\mathbf{a}'\mathbf{X}^{(1)}, \mathbf{b}'\mathbf{X}^{(2)}) = \frac{\mathbf{c}'\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1/2}\mathbf{d}}{\sqrt{\mathbf{c}'\mathbf{c}}\sqrt{\mathbf{d}'\mathbf{d}}} \le \sqrt{\lambda_k}$$

with equality for $\mathbf{c} = \mathbf{e}_k$ or $\mathbf{a} = \boldsymbol{\Sigma}_{11}^{-1/2}\mathbf{e}_k$ and $\mathbf{b} = \boldsymbol{\Sigma}_{22}^{-1/2}\mathbf{f}_k$, as before. Thus, $U_k = \mathbf{e}_k'\boldsymbol{\Sigma}_{11}^{-1/2}\mathbf{X}^{(1)}$ and $V_k = \mathbf{f}_k'\boldsymbol{\Sigma}_{22}^{-1/2}\mathbf{X}^{(2)}$, are the kth canonical pair, and they have correlation $\sqrt{\lambda_k} = \rho_k^*$.

Although we did not explicitly require the V_k to be uncorrelated,

$$\operatorname{Cov}(V_k, V_\ell) = \mathbf{f}_k'\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{\Sigma}_{22}\boldsymbol{\Sigma}_{22}^{-1/2}\mathbf{f}_\ell = \mathbf{f}_k'\mathbf{f}_\ell = 0, \qquad \text{if } k \neq \ell \le p$$

Also,

$$\operatorname{Cov}(U_k, V_\ell) = \mathbf{e}_k'\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1/2}\mathbf{f}_\ell = 0, \qquad \text{if } k \neq \ell \le p$$

since $\mathbf{f}_k'$ is a multiple of $\mathbf{e}_k'\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1/2}$ by (5). ■