Appendix A.1

DEFINITE QUADRATIC FORMS SUBJECT TO CONSTRAINTS

DANIEL McFADDEN

University of California, Berkeley

1. Conditions for a Matrix to be Positive Definite

Let $\mathbf{A}$ denote a real symmetric matrix of order n. Let σ denote a subvector $(i_1,\ldots,i_r)$ of the vector of integers $(1,\ldots,n)$, ordered so that $i_1 < i_2 < \cdots < i_r$. Where the order r of σ is not clear from the context, we write σ_r. Let $\mathbf{S}_r$ denote the set of subvectors σ of order r; $\mathbf{S}_r$ contains ${}_n\mathbf{C}_r$ elements. For $r = 0,\ldots,n-1$, let $\mathbf{A}_{\sigma_r}$ denote the matrix formed from the rows and columns of $\mathbf{A}$ which are not contained in σ_r; i.e., the matrix formed by deleting the rows and columns in σ_r. The determinant $|\mathbf{A}_{\sigma_r}|$ is termed an $(n-r)$th order principal minor of $\mathbf{A}$.

We assume the following basic matrix results to be known:

(1) Associated with the symmetric matrix $\mathbf{A}$ are n (not necessarily distinct) real characteristic values, given by the roots of the polynomial $P(\lambda) = |\mathbf{A} - \lambda\mathbf{I}|$.

(2) The characteristic polynomial has the expansion

$$P(\lambda) = (-\lambda)^n + k_{n-1}(-\lambda)^{n-1} + \cdots + k_1(-\lambda) + k_0,$$

where

$$k_j = \sum_{\sigma_j \in S_j} |\mathbf{A}_{\sigma_j}|$$

is the sum of the $(n-j)$th order principal minors of $\mathbf{A}$.

(3) There exists a matrix $\mathbf{T}$ such that $\mathbf{T'T} = \mathbf{I}$ and $\mathbf{T'AT} = \mathbf{D}$, where $\mathbf{D}$ is a diagonal matrix whose diagonal elements are the characteristic values of $\mathbf{A}$.

(4) The determinant of a matrix equals the product of its characteristic values, and the determinant of a product of matrices equals the product of the determinants.

(5) If $\mathbf{A}$ is positive definite (i.e., $\mathbf{x'x} = 1$ implies $\mathbf{x'Ax} > 0$), then $\mathbf{A}_{\sigma_r}$ is positive definite for each σ_r.

(6) $\mathbf{A}$ is positive definite if and only if all the characteristic values of $\mathbf{A}$ are positive.

The following result relates positive definiteness to properties of the principal minors.

Lemma 1. The following conditions are equivalent:

(i) $\mathbf{A}$ is positive definite.
(ii) All principal minors of $\mathbf{A}$ are positive; i.e., $|\mathbf{A}_\sigma| > 0$ for all $\sigma \in \mathbf{S}_r$ $r = 0,\ldots,n-1$.
(iii) The sum of the principle minors of order $n - r$ is positive for $r = 0,\ldots,n-1$; i.e., $\sum_{s \in S_r} |\mathbf{A}_\sigma| > 0$.
(iv) For each $r = 0,\ldots,n-1$, $|\mathbf{A}_{\sigma_r}| \geqq 0$ and for at least one $\sigma \in \mathbf{S}_r$, $|\mathbf{A}_\sigma| > 0$.
(v) There exists at least one nested sequence of positive principal minors; i.e., there exist $\sigma_0 \subseteq \sigma_1 \subseteq \cdots \subseteq \sigma_{n-1}$ such that $|\mathbf{A}_{\sigma_r}| > 0$.

Proof: The equivalence of (i) and (ii) follows from the basic matrix results (4), (5), and (6). Clearly, (ii) implies (iv) implies (iii) and (ii) implies (v). The proof is completed by showing (iii) implies (i) and (v) implies (i).

Suppose (iii) holds. Then, the coefficient k_{n-j} in the characteristic polynomial $|\mathbf{A} - \lambda \mathbf{I}| = (-\lambda)^n + k_{n-1}(-\lambda)^{n-1} + \cdots + k_0$ equals $\sum_{\sigma \in S_{n-j}} |\mathbf{A}_\sigma|$. Since the characteristic values of $\mathbf{A}$ are all real and the k_{n-j} are all positive, Descarte's rule of signs implies the roots are all positive. Basic result (6) then implies (i).

Suppose (v) holds. We shall show that (i) holds by induction on the order of the matrix n. The result holds trivially for $n = 1$. Suppose that it has been proved for matrices of order up through $n - 1$, and consider a matrix $\mathbf{A}$ of order n with property (v). Assume without loss of generality that the rows and columns of A have been permuted so that the nested sequence of positive minors is formed by successively deleting the first (remaining) row and column of $|\mathbf{A}|$. Write A as a partitioned matrix

$$\mathbf{A} = \left[\begin{array}{c|c} a_{11} & \mathbf{a}' \\ \hline \mathbf{a} & \mathbf{A}_\sigma \end{array}\right],$$

where $\mathbf{a}' = (a_{12},\ldots,a_{1n})$ and $\sigma = (1)$. Expanding $|\mathbf{A}|$ about elements of the first row and the first column yields

$$|\mathbf{A}| = a_{11}|\mathbf{A}_\sigma| + \sum_{i=2}^{n}(-1)^i a_{1i}\sum_{j=2}^{n}(-1)^{j-1}a_{j1}|\mathbf{A}_\sigma^{ij}|,$$

where $|\mathbf{A}_\sigma^{ij}|$ is the minor formed from $|\mathbf{A}_\sigma|$ by deleting row i and column j. Since $|\mathbf{A}_\sigma| \neq 0$, $\mathbf{A}_\sigma^{-1}$ exists and its ijth element is $(\mathbf{A}_\sigma^{-1})_{ij} = (-1)^{i+j}|\mathbf{A}_\sigma^{ij}|/|\mathbf{A}_\sigma|$. Hence,

$$|\mathbf{A}| = a_{11}|\mathbf{A}_\sigma| - |\mathbf{A}_\sigma|\mathbf{a}'\mathbf{A}_\sigma^{-1}\mathbf{a}.$$

Since (v) implies (i) for matrices of order $n-1$, $\mathbf{A}_\sigma$ is positive definite, implying $\mathbf{A}_\sigma^{-1}$ positive definite. Now consider a vector $(x_0,\mathbf{x}')$ with $\mathbf{x}'$ of order $n-1$, $x_0^2 + \mathbf{x}'\mathbf{x} = 1$, and consider the quadratic form

$$Q = (x_0,\mathbf{x}')\mathbf{A}\begin{pmatrix} x_0 \\ \mathbf{x} \end{pmatrix} = x_0^2 a_{11} + 2x_0\mathbf{a}'\mathbf{x} + \mathbf{x}'\mathbf{A}_\sigma\mathbf{x}.$$

Using the expression obtained above to eliminate a_{11},

$$\begin{aligned} Q &= x_0^2|\mathbf{A}|/|\mathbf{A}_\sigma| + x_0^2\mathbf{a}'\mathbf{A}_\sigma^{-1}\mathbf{a} + 2x_0\mathbf{a}'\mathbf{A}_\sigma^{-1}\mathbf{A}_\sigma\mathbf{x} + \mathbf{x}'\mathbf{A}_\sigma\mathbf{A}_\sigma^{-1}\mathbf{A}_\sigma\mathbf{x} \\ &= x_0^2|\mathbf{A}|/|\mathbf{A}_\sigma| + (x_0\mathbf{a} + \mathbf{A}_\sigma\mathbf{x})'\mathbf{A}_\sigma^{-1}(x_0\mathbf{a} + \mathbf{A}_\sigma\mathbf{x}). \end{aligned}$$

Since $\mathbf{A}_\sigma^{-1}$ is positive definite, the second term in Q is non-negative, and is positive if $x_0 = 0$, while the first term is non-negative and is positive if $x_0 \neq 0$. Hence $Q > 0$ and A is positive definite. By induction, (v) implies (i). Q.E.D.

An example shows that it is essential that the positive principal minors in condition (v) of the lemma be nested. The matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

is not positive definite, but has a positive principal minor of each order.

2. Conditions for a Matrix to be Positive Definite Subject to Constraint

Let $\mathbf{A}$ denote a real symmetric matrix of order n, and $\mathbf{B}$ denote a real $m \times n$ matrix of rank m, with $m < n$. We say $\mathbf{A}$ is positive definite subject to constraint $\mathbf{B}$ if $\mathbf{x}'\mathbf{x} = 1$ and $\mathbf{Bx} = \mathbf{0}$ implies $\mathbf{x}'\mathbf{Ax} > 0$.

> *Lemma* 2. $\mathbf{A}$ is positive definite subject to constraint $\mathbf{B}$ if and only if there exists $\lambda_0 \geqq 0$ such that $\mathbf{A} + \lambda \mathbf{B}'\mathbf{B}$ is positive definite for $\lambda \geqq \lambda_0$. (*Note*: This lemma does not require that $\mathbf{B}$ be of full rank.)

Proof: Suppose $\mathbf{A} + \lambda \mathbf{B}'\mathbf{B}$ is positive definite. Then $\mathbf{x}'\mathbf{x} = 1$ implies $\mathbf{x}'(\mathbf{A} + \lambda \mathbf{B}'\mathbf{B})\mathbf{x} > 0$. If $\mathbf{Bx} = \mathbf{0}$, then $\mathbf{x}'\mathbf{Ax} > 0$ and $\mathbf{A}$ is positive definite subject to constraint $\mathbf{B}$.

Suppose there exists a sequence $\lambda_k \to +\infty$ such that $\mathbf{A} + \lambda_k \mathbf{B}'\mathbf{B}$ fails to be positive definite; i.e., there exists $\mathbf{x}_k$ such that $\mathbf{x}_k'\mathbf{x}_k = 1$ and $\mathbf{x}_k'(\mathbf{A} + \lambda_k \mathbf{B}'\mathbf{B})\mathbf{x}_k \leqq 0$. The $\mathbf{x}_k$ have a subsequence converging to $\mathbf{x}_*$. Since $\mathbf{x}_k'\mathbf{Ax}_k$ is bounded, $\mathbf{x}_k'\mathbf{B}'\mathbf{Bx}_k \geqq 0$, and $\lambda_k \to +\infty$, we have $\lim_k \mathbf{Bx}_k = \mathbf{Bx}_* = \mathbf{0}$ and $\mathbf{x}_*'\mathbf{Ax}_* \leqq \limsup -\lambda_k \mathbf{x}_k'\mathbf{B}'\mathbf{Bx}_k \leqq 0$. But $\mathbf{Bx}_* = \mathbf{0}$ and $\mathbf{x}_*'\mathbf{Ax}_* \leqq 0$ imply $\mathbf{A}$ is not positive definite subject to constraint $\mathbf{B}$. Q.E.D.

In examining conditions for $\mathbf{A}$ to be positive definite subject to constraint $\mathbf{B}$, we shall utilize the symmetric $n + m$ bordered matrix

$$\mathbf{H} = \left[\begin{array}{c|c} \mathbf{A} & \mathbf{B}' \\ \hline \mathbf{B} & \mathbf{0} \end{array}\right].$$

The relevance of this matrix is established by the following argument: A necessary and sufficient condition for $\mathbf{A}$ to be positive definite subject to constraint $\mathbf{B}$ is that one have a positive solution to

$$\operatorname{Min}\{\mathbf{x}'\mathbf{Ax} \mid \mathbf{x}'\mathbf{x} = 1 \text{ and } \mathbf{Bx} = \mathbf{0}\}. \tag{1}$$

The following lemma establishes that the solution of (1) can be found by examining the critical points of the Lagrangian

$$\mathbf{L} = \tfrac{1}{2}\mathbf{x}'\mathbf{Ax} + \tfrac{1}{2}\lambda(1 - \mathbf{x}'\mathbf{x}) + \mathbf{p}'\mathbf{Bx}. \tag{2}$$

> *Lemma* 3. If $\mathbf{x}$ solves (1), then there exist $\mathbf{p}$ and λ such that $\lambda = \mathbf{x}'\mathbf{Ax}$, and

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{B}' \\ \hline \mathbf{B} & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \mathbf{x} \\ \hline \mathbf{p} \end{array}\right] = \lambda \left[\begin{array}{c} \mathbf{x} \\ \hline \mathbf{0} \end{array}\right]. \tag{3}$$

Conversely, for any λ solving

$$\left|\begin{array}{c:c} \mathbf{A}-\lambda\mathbf{I} & \mathbf{B}' \\ \hdashline \mathbf{B} & \mathbf{0} \end{array}\right| = 0, \tag{4}$$

there exists $\mathbf{x}$, $\mathbf{p}$ such that $\mathbf{x}'\mathbf{x}=1$, (3) holds, and $\mathbf{x}'\mathbf{A}\mathbf{x}=\lambda$.

Proof: We give a proof of this lemma involving only elementary matrix manipulations. A shorter and more elegant proof could alternately be given by first making an orthogonal change of basis.[1]

We solve (1) by elimination of the constraint. By hypothesis, $\mathbf{B}$ is of rank m and can be partitioned

$$\mathbf{B} = [\mathbf{B}_1 \quad \mathbf{B}_2],$$

where $\mathbf{B}_1$ is $m\times m$ and non-singular and $\mathbf{B}_2$ is $m\times(n-m)$. Partition $\mathbf{x}$ and $\mathbf{A}$ commensurately,

$$\mathbf{x}' = (\mathbf{x}_1' \quad \mathbf{x}_2') \text{ and } \mathbf{A} = \left[\begin{array}{c:c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hdashline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array}\right].$$

If $\mathbf{B}\mathbf{x}=\mathbf{B}_1\mathbf{x}_1+\mathbf{B}_2\mathbf{x}_2=0$, then $\mathbf{x}_1=-\mathbf{B}_1^{-1}\mathbf{B}_2\mathbf{x}_2$ and $\mathbf{x}_1'\mathbf{x}_1+\mathbf{x}_2'\mathbf{x}_2=\mathbf{x}_2'[\mathbf{I}_2+\mathbf{B}_2'(\mathbf{B}_1')^{-1}(\mathbf{B}_1)^{-1}\mathbf{B}_2]\mathbf{x}_2$. Let $\mathbf{C}_2=[\mathbf{I}_2+\mathbf{B}_2'(\mathbf{B}_1\mathbf{B}_1')^{-1}\mathbf{B}_2]^{-1/2}$ and $\mathbf{z}_2=\mathbf{C}_2^{-1}\mathbf{x}_2$. Then $\mathbf{x}'\mathbf{x}=\mathbf{z}_2'\mathbf{z}_2$. Also,

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{z}_2'\mathbf{C}_2[\mathbf{B}_2'(\mathbf{B}_1')^{-1}\mathbf{A}_{11}\mathbf{B}_1^{-1}\mathbf{B}_2 - \mathbf{B}_2'(\mathbf{B}_1')^{-1}\mathbf{A}_{12} - \mathbf{A}_{21}\mathbf{B}_1^{-1}\mathbf{B}_2 + \mathbf{A}_{22}]\mathbf{C}_2\mathbf{z}_2.$$

The minimum of this expression subject to $\mathbf{x}'\mathbf{x}=\mathbf{z}_2'\mathbf{z}_2=1$ is attained for a characteristic vector $\hat{\mathbf{z}}_2$ of this matrix giving its minimum characteristic value λ. Then

$$\mathbf{C}_2[\mathbf{B}_2'(\mathbf{B}_1')^{-1}\mathbf{A}_{11}\mathbf{B}_1^{-1}\mathbf{B}_2 - \mathbf{B}_2'(\mathbf{B}_1')^{-1}\mathbf{A}_{12} - \mathbf{A}_{21}\mathbf{B}_1^{-1}\mathbf{B}_2 + \mathbf{A}_{22}]\mathbf{C}_2\hat{\mathbf{z}}_2 = \lambda\hat{\mathbf{z}}_2,$$

or

$$[\mathbf{B}_2'(\mathbf{B}_1')^{-1}\mathbf{A}_{11}\mathbf{B}_1^{-1}\mathbf{B}_2 - \mathbf{B}_2'(\mathbf{B}_1')^{-1}\mathbf{A}_{12} - \mathbf{A}_{21}\mathbf{B}_1^{-1}\mathbf{B}_2 + \mathbf{A}_{22}]\hat{\mathbf{x}}_2 \\ = \lambda[\mathbf{I}_2 + \mathbf{B}_2'(\mathbf{B}_1\mathbf{B}_1')^{-1}\mathbf{B}_2]\hat{\mathbf{x}}_2,$$

where $\hat{\mathbf{x}}_2=\mathbf{C}_2\hat{\mathbf{z}}_2$. For the above value of λ, we wish to show that there exists a $\mathbf{p}$ such that (3) is satisfied. A solution must satisfy

$$\begin{aligned} \mathbf{A}_{11}\hat{\mathbf{x}}_1 + \mathbf{A}_{12}\hat{\mathbf{x}}_2 + \mathbf{B}_1'\mathbf{p} &= \lambda\hat{\mathbf{x}}_1, \\ \mathbf{A}_{21}\hat{\mathbf{x}}_1 + \mathbf{A}_{22}\hat{\mathbf{x}}_2 + \mathbf{B}_2'\mathbf{p} &= \lambda\hat{\mathbf{x}}_2. \end{aligned} \tag{5}$$

[1]There exists an orthogonal $n\times n$ matrix $\mathbf{S}$ such that $\tilde{\mathbf{B}}=\mathbf{B}\mathbf{S}'=[\mathbf{I}_m : \mathbf{0}_{n-m}]$, where $\mathbf{I}_m$ is the identity matrix and $\mathbf{0}_{n-m}$ is an $m\times(n-m)$ matrix of zeroes. Let $\tilde{\mathbf{A}}=\mathbf{S}\mathbf{A}\mathbf{S}'$. Then, it is straightforward to establish Lemma 3 for the problem of minimizing $\mathbf{x}'\tilde{\mathbf{A}}\mathbf{x}$ subject to $\mathbf{x}'\mathbf{x}=1$ and $\tilde{\mathbf{B}}\mathbf{x}=0$, and then to show that (3) and (4) are invariant under an orthogonal change of bias. This argument was suggested by S. Cosslett.

Solving for $\mathbf{p}$ from the first set of equations yields

$$\begin{aligned}\mathbf{p} &= (\mathbf{B}_1')^{-1}[\lambda\hat{\mathbf{x}}_1 - \mathbf{A}_{11}\hat{\mathbf{x}}_1 - \mathbf{A}_{12}\hat{\mathbf{x}}_2] \\ &= (\mathbf{B}_1')^{-1}[-\lambda\mathbf{B}_1^{-1}\mathbf{B}_2\hat{\mathbf{x}}_2 + \mathbf{A}_{11}\mathbf{B}_1^{-1}\mathbf{B}_2\hat{\mathbf{x}}_2 - \mathbf{A}_{12}\hat{\mathbf{x}}_2].\end{aligned}$$

Substituting this expression into the second set of equations,

$$\begin{aligned}\mathbf{A}_{21}\hat{\mathbf{x}}_1 + \mathbf{A}_{22}\hat{\mathbf{x}}_2 + \mathbf{B}_2'\mathbf{p} - \lambda\hat{\mathbf{x}}_2 &= -\mathbf{A}_{21}\mathbf{B}_1^{-1}\mathbf{B}_2\hat{\mathbf{x}}_2 + \mathbf{A}_{22}\hat{\mathbf{x}}_2 - \lambda\mathbf{B}_2'(\mathbf{B}_1')^{-1}\mathbf{B}_1^{-1}\mathbf{B}_2\hat{\mathbf{x}}_2 \\ &\quad + \mathbf{B}_2'(\mathbf{B}_1')^{-1}\mathbf{A}_{11}\mathbf{B}_1^{-1}\mathbf{B}_2\hat{\mathbf{x}}_2 - \mathbf{B}_2'(\mathbf{B}_1')^{-1}\mathbf{A}_{12}\hat{\mathbf{x}}_2 - \lambda\hat{\mathbf{x}}_2 \\ &= -\lambda[\mathbf{I} + \mathbf{B}_2'(\mathbf{B}_1\mathbf{B}_1')^{-1}\mathbf{B}_2]\hat{\mathbf{x}}_2 \\ &\quad + [\mathbf{B}_2'(\mathbf{B}_1')^{-1}\mathbf{A}_{11}\mathbf{B}_1^{-1}\mathbf{B}_2 - \mathbf{A}_{21}\mathbf{B}_1^{-1}\mathbf{B}_2 \\ &\quad - \mathbf{B}_2'(\mathbf{B}_1')^{-1}\mathbf{A}_{12} + \mathbf{A}_{22}]\hat{\mathbf{x}}_2 = \mathbf{0},\end{aligned}$$

by the characteristic value property. Hence, the system (5) has a consistent solution for $\mathbf{p}$. Premultiplying the last equation by $\hat{\mathbf{x}}_2$, and noting that $\hat{\mathbf{x}}_2'[\mathbf{I} + \mathbf{B}_2'(\mathbf{B}_1'\mathbf{B}_1)^{-1}\mathbf{B}_2]\hat{\mathbf{x}}_2 = 1$ and $\hat{\mathbf{x}}_1 = -\mathbf{B}_1^{-1}\mathbf{B}_2\hat{\mathbf{x}}_2$, we obtain $\lambda = \hat{\mathbf{x}}'\mathbf{A}\hat{\mathbf{x}}$. Hence, (3) holds.

Suppose λ is a root of the polynomial (4). Then equation (3) holds for some $\mathbf{x}$ and $\mathbf{p}$ not both identically zero. If $\mathbf{x} = \mathbf{0}$, one has $\mathbf{B}'\mathbf{p} = \mathbf{0}$, contradicting the assumption that $\mathbf{B}$ is of rank m. Hence $\mathbf{x} \neq \mathbf{0}$ and we can normalize $\mathbf{x}'\mathbf{x} = 1$. Then $\mathbf{x}'(\mathbf{A}\mathbf{x} + \mathbf{B}'\mathbf{p}) = \lambda\mathbf{x}'\mathbf{x}$ and $\mathbf{B}\mathbf{x} = \mathbf{0}$ implies $\mathbf{x}'\mathbf{A}\mathbf{x} = \lambda\mathbf{x}'\mathbf{x} = \lambda$. Q.E.D.

As before, let σ be a subvector of $(1,\ldots,n)$, and $\mathbf{A}_\sigma$ be the matrix formed from the rows and columns of $\mathbf{A}$ not contained in σ. For the $m \times n$ matrix $\mathbf{B}$, let $\mathbf{B}_\sigma$ be the submatrix formed by deleting columns from σ. Hence, if $\sigma = \sigma_r$ contains r elements, then $\mathbf{A}_\sigma$ is a square matrix of order $n - r$ and $\mathbf{B}_\sigma$ is $m \times (n - r)$. The following result relates positive definiteness subject to constraint to properties of the principal minors.

Lemma 4. If $\mathbf{A}$ is an $n \times n$ symmetric matrix and $\mathbf{B}$ is an $m \times n$ matrix of rank $m < n$, then the following conditions are equivalent:

(i) $\mathbf{A}$ is positive definite subject to constraint $\mathbf{B}$.

(ii) For $r = 0,\ldots,n - m$,

$$(-1)^m \sum_{\sigma \in \mathbf{S}_r} \begin{vmatrix} \mathbf{A}_\sigma & \mathbf{B}_\sigma' \\ \mathbf{B}_\sigma & \mathbf{0} \end{vmatrix} > 0.$$

(iii) For $r = 0,\ldots,n - m$, and $\sigma \in \mathbf{S}_r$,

$$(-1)^m \begin{vmatrix} \mathbf{A}_\sigma & \mathbf{B}_\sigma' \\ \mathbf{B}_\sigma & \mathbf{0} \end{vmatrix} \geqq 0,$$

with strict inequality holding for at least one $\sigma \in \mathbf{S}_r$.

(iv) There exists at least one nested sequence of principle minors of order $2m$ through $n+m$ formed by deleting symmetric rows and columns from the first n rows and columns, which have the sign of $(-1)^m$; i.e., there exist $\sigma_0 \subseteq \sigma_1 \subseteq \cdots \subseteq \sigma_{n-m}$ such that

$$(-1)^m \begin{vmatrix} \mathbf{A}_{\sigma_r} & \mathbf{B}'_{\sigma_r} \\ \mathbf{B}_{\sigma_r} & \mathbf{0} \end{vmatrix} > 0.$$

Further, note that (iv) implies $|\mathbf{B}_{\sigma_{n-m}}| \neq 0$, while if (i) holds and $|\mathbf{B}_{\sigma_{n-m}}| \neq 0$ for some σ_{n-m}, then any nested sequence starting from σ_{n-m} satisfies (iv).

Proof: We first show that (i) and (ii) are equivalent. By Lemma 3, (i) holds if and only if the roots of the polynomial in (4) are all positive. Expand this polynomial in powers of $(-\lambda)$,

$$\begin{vmatrix} \mathbf{A}-\lambda\mathbf{I} & \mathbf{B}' \\ \mathbf{B} & \mathbf{0} \end{vmatrix} = \sum_{j=0}^{n+m} k_j(-\lambda)^j. \tag{6}$$

The argument below establishes that

$$k_j = 0 \quad \text{for} \quad j > n-m,$$

and

$$k_j = \sum_{\sigma \in \mathbf{S}_j} \begin{vmatrix} \mathbf{A}_{\sigma} & \mathbf{B}'_{\sigma} \\ \mathbf{B}_{\sigma} & \mathbf{0} \end{vmatrix} \quad \text{for} \quad j \leqq n-m.$$

A term in a full expansion of the determinant (6) is the product of $n-l$ elements from the northwest submatrix, l elements from each of the northeast and southwest submatrices, and $m-l$ elements from the southeast submatrix. This term can be non-zero only if $m-l=0$, implying at most $n-m$ elements are taken from the northwest submatrix. Hence, $k_j = 0$ for $j > n-m$. Consider the collection of all terms in a full expansion of (6) which contribute to k_j for some $j \leqq n-m$. Each such term is the product of factors $(-\lambda)$ taken from diagonal elements of the northwest submatrix in (6), corresponding to columns in σ_j; multiplied by a term in the expansion of the determinant formed by deleting the σ_j rows and columns from (6); i.e., a term in the expansion of

$$\begin{vmatrix} \mathbf{A}_{\sigma_j} & \mathbf{B}'_{\sigma_j} \\ \mathbf{B}_{\sigma_j} & \mathbf{0} \end{vmatrix}.$$

Then, k_j equals the sum over $\sigma \in \mathbf{S}_j$ of terms of the form

$$\begin{vmatrix} \mathbf{A}_{\sigma} & \mathbf{B}'_{\sigma} \\ \mathbf{B}_{\sigma} & \mathbf{0} \end{vmatrix},$$

as was to be demonstrated.

For $\sigma \in \mathbf{S}_{n-m}$,

$$\left|\begin{array}{c:c} \mathbf{A}_\sigma & \mathbf{B}'_\sigma \\ \hdashline \mathbf{B}_\sigma & \mathbf{0} \end{array}\right| = (-1)^{m^2} \left|\begin{array}{c:c} \mathbf{B}'_\sigma & \mathbf{A}_\sigma \\ \hdashline \mathbf{0} & \mathbf{B}_\sigma \end{array}\right| = (-1)^m |\mathbf{B}_\sigma|^2.$$

Since $\mathbf{B}$ is of rank m, $|\mathbf{B}_\sigma| \neq 0$ for some $\sigma \in \mathbf{S}_{n-m}$, implying $(-1)^m k_{n-m} > 0$. By Descartes' rule of signs, the roots of the polynomial (6), which are real since the matrix is symmetric, are all positive if and only if the coefficients $k_0,\ldots,k_{n-m}$ are of uniform sign. Hence, (i) and (ii) are equivalent.

It is trivial that (iii) implies (ii). We next show that (i) implies (iii). If A is positive definite subject to constraint B, then for any $\sigma \in \mathbf{S}_r$, $r = 0,\ldots,n-m$, $\mathbf{A}_\sigma$ is positive definite subject to constraint $\mathbf{B}_\sigma$. (This can be seen by setting the components of $\mathbf{x}$ outside of σ equal to zero in the definition $\mathbf{x}'\mathbf{x} = 1$, $\mathbf{Bx} = \mathbf{0}$ implies $\mathbf{x}'\mathbf{Ax} > 0$.) By Lemma 2, $\mathbf{A}_\sigma + \lambda \mathbf{B}'_\sigma \mathbf{B}_\sigma$ is positive definite for λ sufficiently large. Further, if $\mathbf{B}_\sigma$ is of rank m, then by Lemma 3,

$$\left|\begin{array}{c:c} \mathbf{A}_\sigma & \mathbf{B}'_\sigma \\ \hdashline \mathbf{B}_\sigma & \mathbf{0} \end{array}\right| \neq 0.$$

Consider the matrices

$$\mathbf{C}_\sigma = \left|\begin{array}{c:c} \mathbf{A}_\sigma & \lambda \mathbf{B}'_\sigma \\ \hdashline \mathbf{B}_\sigma & -\mathbf{I}_m \end{array}\right| \text{ and } \mathbf{E}_\sigma = \left|\begin{array}{c:c} \mathbf{I}_\sigma & \mathbf{0} \\ \hdashline \mathbf{B}_\sigma & \mathbf{I}_m \end{array}\right|.$$

Then

$$\mathbf{C}_\sigma \mathbf{E}_\sigma = \left|\begin{array}{c:c} \mathbf{A}_\sigma + \lambda \mathbf{B}'_\sigma \mathbf{B}_\sigma & \lambda \mathbf{B}'_\sigma \\ \hdashline \mathbf{0} & -\mathbf{I}_m \end{array}\right|,$$

$$|\mathbf{E}_\sigma| = 1 \quad \text{and} \quad |\mathbf{C}_\sigma|\,|\mathbf{E}_\sigma| = |\mathbf{C}_\sigma \mathbf{E}_\sigma| = (-1)^m |\mathbf{A}_\sigma + \lambda \mathbf{B}'_\sigma \mathbf{B}_\sigma|.$$

Consider the limit

$$\lim_{\lambda\to+\infty} \lambda^{-m}(-1)^m |\mathbf{C}_\sigma| = \lim_{\lambda\to+\infty} (-1)^m \left|\begin{array}{c:c} \mathbf{A}_\sigma & \mathbf{B}'_\sigma \\ \hdashline \mathbf{B}_\sigma & -(1/\lambda)\mathbf{I}_m \end{array}\right| = (-1)^m \left|\begin{array}{c:c} \mathbf{A}_\sigma & \mathbf{B}'_\sigma \\ \hdashline \mathbf{B}_\sigma & \mathbf{0} \end{array}\right|.$$

Since $(-1)^m |\mathbf{C}_\sigma| > 0$, one obtains in the limit

$$(-1)^m \left|\begin{array}{c:c} \mathbf{A}_\sigma & \mathbf{B}'_\sigma \\ \hdashline \mathbf{B}_\sigma & \mathbf{0} \end{array}\right| \geqq 0. \tag{7}$$

Further, if $\mathbf{B}_\sigma$ is of rank m, the inequality in (7) is strict. Since $\mathbf{B}_\sigma$ must be of full rank for some $\sigma \in \mathbf{S}_r$, this establishes that (iii) holds.

Next we show that (i) and (iv) are equivalent. Suppose (iv) holds.

Then for $\sigma = \sigma_{n-m}$,

$$0 < (-1)^m \begin{vmatrix} \mathbf{A}_\sigma & \mathbf{B}'_\sigma \\ \mathbf{B}_\sigma & \mathbf{0} \end{vmatrix} = |\mathbf{B}_\sigma|^2,$$

implying $\mathbf{B}'_\sigma\mathbf{B}_\sigma$ positive definite. Hence, for λ sufficiently large, $\mathbf{A}_{\sigma_{n-m}} + \lambda\mathbf{B}'_{\sigma_{n-m}}\mathbf{B}_{\sigma_{n-m}}$ is positive definite, implying the existence of a nested sequence of positive principle minors of $\mathbf{A}_{\sigma_{n-m}} + \mathbf{B}'_{\sigma_{n-m}}\mathbf{B}_{\sigma_{n-m}}$. Further, (iv) implies the existence of a nested sequence $\sigma_0 \subseteq \cdots \subseteq \sigma_{n-m}$, such that for $\sigma = \sigma_r$, $r = 0,\ldots,n-m$,

$$(-1)^m \begin{vmatrix} \mathbf{A}_\sigma & \mathbf{B}'_\sigma \\ \mathbf{B}_\sigma & \mathbf{0} \end{vmatrix} > 0. \tag{8}$$

Since $\lim_{\lambda\to\infty}(-\lambda)^{-m}|\mathbf{C}_\sigma|$ equals the left-hand side of (8), we conclude that for λ sufficiently large, $(-1)^m|\mathbf{C}_\sigma| = |\mathbf{A}_\sigma + \lambda\mathbf{B}'_\sigma\mathbf{B}_\sigma|$ is positive for $r = 0,\ldots,n-m$. Hence, since $(\mathbf{A}+\lambda\mathbf{B}'\mathbf{B})_\sigma = \mathbf{A}_\sigma + \lambda\mathbf{B}'_\sigma\mathbf{B}_\sigma$, we have established the existence of a full nested sequence of positive principal minors of $\mathbf{A}+\lambda\mathbf{B}'\mathbf{B}$. By Lemma 1, $\mathbf{A}+\lambda\mathbf{B}'\mathbf{B}$ is then positive definite, and by Lemma 2, condition (i) holds.

Next suppose (i) holds, and choose σ_{n-m} such that $|\mathbf{B}_{\sigma_{n-m}}| \neq 0$. (This can be done since $\mathbf{B}$ is of rank m.) Choose any σ_r such that $\sigma_0 \subseteq \sigma_1 \subseteq \cdots \subseteq \sigma_{n-m}$ for $r = 0,\ldots,n-m-1$. The proof above that (i) implies (iii) establishes that

$$(-1)^m \begin{vmatrix} \mathbf{A}_{\sigma_r} & \mathbf{B}'_{\sigma_r} \\ \mathbf{B}_{\sigma_r} & \mathbf{0} \end{vmatrix} > 0,$$

for $r = 0,\ldots,n-m$. Hence, (iv) holds.

The note at the end of the lemma follows from the fact that

$$(-1)^m \begin{vmatrix} \mathbf{A}_{\sigma_{n-m}} & \mathbf{B}'_{\sigma_{n-m}} \\ \mathbf{B}_{\sigma_{n-m}} & \mathbf{0} \end{vmatrix} = |\mathbf{B}_{\sigma_{n-m}}|^2,$$

and from the construction used to establish that (i) implies (iii) and (iv). Q.E.D.