Caltech Division of the Humanities and Social Sciences

More than you wanted to know about quadratic forms

KC Border

v. 2018.11.11::17.30

Contents

1 Quadratic forms

In these notes, I shall treat vectors as column matrices, and use $'$ to denote matrix transposition. Thus the Euclidean inner product $x \cdot y = \sum_{i=1}^{n} x_i y_i$ of n-vectors can also be written as $x'y$.

Let A be an $n \times n$ symmetric matrix, and let x be an n-vector. Then $x \cdot Ax = x'Ax$ is a scalar,

$$x'Ax = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j. \tag{1}$$

The mapping $Q\colon x \mapsto x'Ax$ is the **quadratic form** defined by A.[1]

A symmetric matrix A (or its associated quadratic form) is called

- **positive definite** if $x'Ax > 0$ for all nonzero x.
- **negative definite** if $x'Ax < 0$ for all nonzero x.
- **positive semidefinite** if $x'Ax \geqslant 0$ for all x.
- **negative semidefinite** if $x'Ax \leqslant 0$ for all x.

The reason for restricting attention to symmetric matrices is so that their eigenvectors generate an orthonormal basis for $\boldsymbol{R}^{\mathrm{n}}$. If A is not symmetric, then $\frac{A+A'}{2}$ is symmetric and $x'Ax = x \cdot (\frac{A+A'}{2})x$ for any x, so we do not lose much applicability by this assumption. Some authors use the term **quasi (semi)definite** when they do not wish to impose symmetry.

Proposition 1 (Eigenvalues and definiteness) *The symmetric matrix A is*

1. *positive definite if and only if all its eigenvalues are strictly positive.*
2. *negative definite if and only if all its eigenvalues are strictly negative.*
3. *positive semidefinite if and only if all its eigenvalues are nonnegative.*
4. *negative semidefinite if and only if all its eigenvalues are nonpositive.*

Proof: Let $\{x^1, \ldots, x^n\}$ be an orthonormal basis for $\boldsymbol{R}^{\mathrm{n}}$ consisting of eigenvectors of A. (See, e.g., Apostol [2, Theorem 5.4, p. 120].) Let λ_i be the eigenvalue corresponding to x^i. That is,

$$Ax^i = \lambda_i x^i.$$

[1] For decades I was baffled by the term *form.* I once asked Tom Apostol at a faculty cocktail party what it meant. He professed not to know (it was a cocktail party, so that is excusable), but suggested that I should ask John Todd. He hypothesized that mathematicians don't know the difference between form and function, a clever reference to modern architectural philosophy. I was too intimidated by Todd to ask, but I subsequently learned (where, I can't recall) that *form* refers to a polynomial function in several variables where each term in the polynomial has the same degree. (The *degree* of the term is the sum of the exponents. For example, in the expression $xyz + x^2y + xz + z$, the first two terms have degree three, the third term has degree two and the last one has degree one. It is thus not a form.) This is most often encountered in the phrases *linear form* (each term has degree one) or *quadratic form* (each term has degree two).

Writing $y = \sum_{j=1}^{n}(\alpha_j x^y)$, we have

$$Ay = A\Big(\sum_{j=1}^{n} \alpha_j x^i\Big) = \sum_{j=1}^{n} \alpha_j A x^j = \sum_{j=1}^{n} \alpha_j \lambda_j x^j,$$

so

$$y \cdot Ay = \Big(\sum_{i=1}^{n} \alpha_i x^i\Big) \cdot \Big(\sum_{j=1}^{n} \alpha_j \lambda_j x^j\Big) = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j \lambda_j x^i \cdot x^j = \sum_{k=1}^{n} (\alpha_k)^2 \lambda_k,$$

where the last equality follows from the orthonormality of $\{x^1, \ldots, x^n\}$. All the statements above follow from this equation and the fact that $(\alpha_k)^2 \geqslant 0$ for all k. ∎

Proposition 2 (Definiteness of the inverse) *If A is positive definite (negative definite), then A^{-1} exists and is also positive definite (negative definite).*

Proof: First off, how do we know the inverse of A exists? Suppose $Ax = 0$. Then $x'Ax = x'0 = 0$. Since A is positive definite, we see that $x = 0$. Therefore A is invertible. Here are two proofs of the proposition.

First proof. Since $(Ax = \lambda x) \implies (x = \lambda A^{-1}x) \implies (A^{-1}x = \frac{1}{\lambda}x)$, the eigenvalues of A and A^{-1} are reciprocals, so they must have the same sign. Apply Proposition 1.

Second proof.

$$x'A^{-1}x = y \cdot Ay \quad \text{where} \quad y = A^{-1}x.$$

∎

We now digress a bit. Recall that the **characteristic polynomial** f of a square matrix A is defined by $f(\lambda) = \det(\lambda I - A)$. Roots of this polynomial are called **characteristic roots** of A.

Lemma 3 *Every eigenvalue of a matrix is a characteristic root, and every real characteristic root is an eigenvalue.*

Proof: To see this note that if λ is an eigenvalue with eigenvector $x \neq 0$, then $(\lambda I - A)x = \lambda x - Ax = 0$, so $(\lambda I - A)$ is singular, so $\det(\lambda I - A) = 0$. That is, λ is a characteristic root of A.

Conversely, if $\det(\lambda I - A) = 0$, then the matrix $\lambda I - A$ is singular, so there is some nonzero vector x with $(\lambda I - A)x = 0$. In other words, $Ax = \lambda x$. ∎

Lemma 4 *The determinant of a square matrix is the product of its characteristic roots.*

Proof: (cf. Apostol [2, p. 106]) Let A be an $n \times n$ square matrix and let f be its characteristic polynomial. Then $f(0) = \det(-A) = (-1)^n \det A$. On the other hand, we can factor f as

$$f(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n)$$

where $\lambda_1, \ldots, \lambda_n$ are its characteristic roots. Thus $f(0) = (-1)^n \lambda_1 \cdots \lambda_n$. ∎

The proof of the next theorem may be found in Debreu [5] or Gantmacher [8, pp. 306–308].

Theorem 5 *For a symmetric matrix A:*

1. *A is positive definite if and only if all its NW principal minors are strictly positive.*
2. *A is negative definite if and only if all its k^{th}-order NW principal minors have sign $(-1)^k$.*
3. *A is positive semidefinite if and only if all its principal minors are nonnegative.*
4. *A is negative semidefinite if and only if all its k^{th}-order principal minors have sign $(-1)^k$ or 0.*

Proof: We start with the necessity of the conditions on the minors.

First note that every principal submatrix of a matrix A inherits its definiteness. To see this let $I \subset \{1, \ldots, n\}$ be the (nonempty) set of indices of rows and columns for the submatrix. Let x be any nonzero vector with $x_k = 0$ for $k \notin I$. Then

$$x'Ax = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j = \sum_{i \in I} \sum_{j \in I} a_{ij} x_i x_j,$$

so the quadratic form defined by the submatrix cannot have a different sign from the quadratic form defined by A.

By Proposition 1, if a matrix is positive definite, all its eigenvalues are positive, so by Lemma 4 its determinant must be positive, as the product of the eigenvalues. Thus every principal submatrix of a positive definite matrix has a strictly positive determinant. Similarly, every principal submatrix of a positive semidefinite matrix has a nonnegative determinant.

The results for negative (semi)definiteness stem from the observation that a matrix A is negative (semi)definite if and only if $-A$ is positive (semi)definite, and that the determinant of a k^{th} order submatrix of $-A$ is $(-1)^k$ times the corresponding subdeterminant of A.

The sufficiency part is harder. To see why such a result might be true, consider first the case $n = 2$. Then, completing the square, we get

$$\begin{aligned} x'Ax &= a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 \\ &= a_{11}\Big(x_1 + \frac{a_{12}}{a_{11}}\,x_2\Big)^2 + \frac{a_{11}a_{22} - a_{12}^2}{a_{11}}\,x_2^2 \\ &= D_1y_1^2 + \frac{D_2}{D_1}\,y_2^2, \end{aligned}$$

where

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & \frac{a_{12}}{a_{11}} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

$D_1 = a_{11}$, the determinant of the 1×1 NW principal minor of A, and $D_2 = \det A$, the determinant of the 2×2 NW principal minor. In this case it is easy to see that $D_1 > 0$ and $D_2 > 0$ imply that A is positive definite.

Lagrange noticed that this technique could be generalized. That is, if $D_1 \neq 0, \dots, D_n \neq 0$ there is always a nonsingular upper triangular matrix U (with 1's on the main diagonal), so that

$$x'Ax = \sum_{i=1}^{n} \frac{D_i}{D_{i-1}} y_i^2,$$

where $y = Ux$, $D_0 = 1$, and D_i is the determinant of the $i \times i$ NW principal minor of A. Given this decomposition, known as Jacobi's formula, it is easy to see why the conditions $D_1 > 0, \dots, D_n > 0$ guarantee that A is positive definite. The matrix U is computed by using Gaussian elimination on A. For details, see, e.g., Gantmacher [8, pp. 33–41, 300–302]. This proves parts (1) and (2).

To prove parts (3) and (4), we use the fact that if A has rank k, then there is a permutation matrix P so that $\hat{A} = P'AP$ satisfies $\hat{D}_1 > 0, \dots, \hat{D}_k > 0$ and $\hat{D}_{k+1} = \cdots = \hat{D}_n = 0$. Furthermore, each $\hat{D}_i$ is some $i \times i$ minor subdeterminant of the original A. Thus there is an upper triangular matrix $\hat{U}$ such that

$$x'Ax = x \cdot P\hat{A}P'x = P'x \cdot \hat{A}P'x = \sum_{i=1}^{k} \frac{\hat{D}_i}{\hat{D}_{i-1}} y_i^2,$$

where $y = \hat{U}P'x$. Again see Gantmacher [8, pp. 33—41] for details. ∎

1.1 Quadratic forms on the unit sphere

In this section we deduce some properties of quadratic forms restricted to subsets of the unit sphere. Consider an $n \times n$ symmetric matrix A. The quadratic form $Q(x) = x'Ax$ is a continuous function of x, so it achieves a maximum on the unit sphere $S = \{x \in \boldsymbol{R}^{\mathrm{n}} : x \cdot x = 1\}$, which is compact. This maximizer turns out to be an eigenvector of A, and the value of the maximum is its corresponding eigenvalue. This eigenvalue also turns out to be the Lagrange Multiplier for the constraint that the maximizer lies on the sphere. We can say even more, for if we restrict attention to the subspace orthogonal to the eigenvector and look for a maximizer, we get another eigenvector and eigenvalue. We can repeat this procedure until we have found them all.

The next proposition is, I believe, well known, but I had trouble finding it written out. It is implicit in Carathéodory [3, § 195].[2] It also follows from result 1f.2.iii in Rao [13, p. 62]. Anderson [1, pp. 273–275] and Franklin [7, Section 6.2, pp. 141–145] give somewhat more explicit treatments.

Proposition 6 (Extrema of quadratic forms on the sphere) *Let A be an $n \times n$ symmetric matrix. Define $x^1, \dots, x^n$ recursively so that x^{k+1} maximizes the quadratic form $Q(x) = x'Ax$ over $S_k = S \cap M_{k\perp}$, where S is the unit sphere in $\boldsymbol{R}^{\mathrm{n}}$, and M_k denotes the span of $x^1, \dots, x^k$, with $M_0 = \{0\}$. Then each x^k, $k = 1, \dots, n$ is an eigenvector of A, and $\lambda_k = Q(x^k)$ is its corresponding eigenvalue.*

Note that by construction $S_{k+1} \subset S_k$, so $\lambda_1 \geqslant \cdots \geqslant \lambda_n$. If A is positive definite, then $\lambda_n > 0$, so we have an alternate proof of Proposition 1.

[2] Be advised that [3] uses the peculiar convention that an expression like $a_{ij}x_j$, where a subscript is repeated, means to sum over that subscript, that is, $a_{ij}x_j$ means $\sum_j a_{ij}x_j$ and $a_{ij}x_ix_j$ means $\sum_i \sum_j a_{ij}x_ix_j$.

Proof[3]: The quadratic form $Q(x) = x'Ax$ is continuously differentiable and $\nabla Q(x) = 2Ax$. Let x^1 be a maximizer of Q on $S = S_0$. (The maximizer is not unique as $((Q(x) = Q(-x).)$ Then x^1 maximizes Q subject to the constraint $1 - x \cdot x = 0$. Now the gradient of this constraint function is $-2x$, which is clearly nonzero (hence linearly independent) on S. It is a nuisance to have these 2s popping up, so let us agree to maximize $\frac{1}{2}x'Ax$ subject $\frac{1}{2}(1 - x \cdot x) = 0$ instead. Therefore by the Lagrange Multiplier Theorem, there exists λ_1 satisfying

$$Ax^1 - \lambda_1 x^1 = 0. \tag{2}$$

This obviously implies that the Lagrange multiplier λ_1 is an eigenvalue of A and x^1 is a corresponding eigenvector. Further, it is the value of the maximum:

$$Q(x^1) = x^1 \cdot Ax^1 = \lambda_1 x^1 \cdot x^1 = \lambda_1, \tag{3}$$

since $x^1 \cdot x^1 = 1$.

We now proceed by induction on k. Let $x^1, \dots, x^k$ be recursively defined as above and assume they satisfy the conclusion of the theorem. Let x^{k+1} be a maximizer of $\frac{1}{2}Q$ over S_k. We wish to show that x^{k+1} is an eigenvector of A and $\lambda_{k+1} = Q(x^{k+1})$ is its corresponding eigenvalue.

By hypothesis, x^{k+1} maximizes $\frac{1}{2}Q(x)$ subject to the constraints

$$\tfrac{1}{2}(1 - x \cdot x) = 0, \quad x \cdot x^1 = 0, \quad \dots \quad x \cdot x^k = 0.$$

The gradients of these constraint functions are $-x$ and $x^1, \dots, x^k$ respectively. By construction, $x^1, \dots, x^{k+1}$ are orthonormal, so at $x = x^{k+1}$ the constraint gradients are linearly independent. Therefore by the Lagrange Multiplier Theorem there exist multipliers λ_{k+1} and $\mu_1, \dots, \mu_k$ satisfying

$$Ax^{k+1} - \lambda_{k+1}x^{k+1} + \mu_1 x^1 + \cdots + \mu_k x^k = 0. \tag{4}$$

Therefore

$$\begin{aligned} Q(x^{k+1}) &= x^{k+1} \cdot Ax^{k+1} \\ &= \lambda_{k+1}x^{k+1} \cdot x^{k+1} - \mu_1 x^{k+1} \cdot x^1 - \cdots - \mu_k x^{k+1} \cdot x^k \\ &= \lambda_{k+1}, \end{aligned} \tag{5}$$

since $x^1, \dots, x^{k+1}$ are orthonormal. That is, the multiplier λ_{k+1} is the maximum value of Q over S_k.

[3]I thank Professor Tony E. Smith, Department of Electrical and Systems Engineering, University of Pennsylvania for pointing out how to simply the proof.

Next note that if $x \in M_{k\perp}$, then $Ax \in M_{k\perp}$. To see this, note that by the induction hypothesis, $Ax_i = \lambda_i x_i$ for $i = 1, \dots, k$. Then since A is symmetric,

$$x_i \cdot Ax = x \cdot Ax_i = x \cdot \lambda_i x_i = 0, \quad i = 1, \dots, k.$$

That is, $Ax \in M_{k\perp}$. By construction $x^{k+1} \in M_{k\perp}$, so we just showed that $Ax^{k+1} \in M_{k\perp}$.

Recall that if $x \perp y$ and $x + y = 0$, then $x = 0$ and $y = 0$.[4] Therefore by equation (4)

$$\underbrace{Ax^{k+1} - \lambda_{k+1}x^{k+1}}_{\in M_{k\perp}} = 0 \quad \text{and} \quad \underbrace{\mu_1 x^1 + \cdots + \mu_k x^k}_{\in M_k} = 0.$$

We conclude therefore that $Ax^{k+1} = \lambda_{k+1}x^{k+1}$, so that x^{k+1} is an eigenvector of A and λ_{k+1} is the corresponding eigenvalue. ∎

2 Quadratic forms under constraint

In subsection 1.1 above, we were interested in a quadratic form restricted to a subspace orthogonal to a set of eigenvectors. In this section we will generalize this problem.

Definition 7 *A matrix A is* ***positive definite under the orthogonality constraints $b^1, \dots, b^m$*** *if it is symmetric and*

$$x'Ax > 0 \quad \text{for all} \quad x \neq 0 \quad \text{satisfying} \quad b^i \cdot x = 0, \quad i = 1, \dots, m.$$

For brevity, when the vectors $b^1, \dots, b^m$ are understood, we often say simply that A is ***positive definite under constraint****. The notions of* ***negative definiteness*** *and* ***semidefiniteness under constraint*** *are defined in the obvious way.*

We can replace $b^1, \dots, b^m$ by any basis for the span of $b^1, \dots, b^m$, so without loss of generality we may assume that $b^1, \dots, b^m$ are linearly independent, or even orthonormal.

[4] This follows from $(x + y) \cdot (x + y) = x \cdot x + 2x \cdot y + y \cdot y = x \cdot x + y \cdot y$ when $x \cdot y = 0$.

2.1 Rank and Definiteness of Quadratic Forms under Constraint

Theorem 8 *Suppose A is an $n \times n$ symmetric matrix that is negative definite under orthogonality constraints for the linearly independent vectors $b^1, \dots, b^m$. That is,*

$$x'Ax < 0 \text{ for all nonzero } x \text{ satisfying } B'x = 0, \tag{6}$$

where B is the $n \times m$ matrix whose j^{th} column is b^j. Then:

1. *The* ***bordered matrix***

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

   is invertible.

2. *Write the* ***inverse bordered matrix*** *as*

$$\left[\begin{array}{c|c} A & B \\ \hline B' & 0 \end{array}\right]^{-1} = \left[\begin{array}{c|c} C & D \\ \hline D' & E \end{array}\right].$$

   Then C is negative semidefinite of rank $n-m$, with $Cx = 0$ if and only if x is a linear combination of $b^1, \dots, b^m$.

Proof: (cf. Samuelson [14, pp. 378–379], Quirk [12, pp. 22–25], and Diewert and Woodland [6, Appendix, Lemma 3])

(1) To see that the bordered matrix is invertible, suppose

$$\left[\begin{array}{c|c} A & B \\ \hline B' & 0 \end{array}\right] \left[\begin{array}{c} x \\ \hline z \end{array}\right] = 0. \tag{7}$$

We wish to show that $x = 0$ and $z = 0$. Now (7) reduces to

$$Ax + Bz = 0 \tag{8}$$

and

$$B'x = 0. \tag{9}$$

Simple calculation shows that the quadratic form associated with the bordered matrix satisfies

$$\left[\begin{array}{c|c} x' & z' \end{array}\right] \left[\begin{array}{c|c} A & B \\ \hline B' & 0 \end{array}\right] \left[\begin{array}{c} x \\ \hline z \end{array}\right] = \left[\begin{array}{c|c} x' & z' \end{array}\right] \left[\begin{array}{c} Ax + Bz \\ \hline B'x \end{array}\right] = x'Ax + 2z'B'x.$$

Now if (7) holds, this quadratic form is zero, so $x'Ax+2z'B'x = 0$. But by (9), this implies $x'Ax = 0$. By assumption, A is definite under the constraint (9) we must have that $x = 0$. Thus 8 implies $Bz = 0$. Since B has linearly independent columns, this implies $z = 0$.

We have just shown that

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

Therefore the bordered matrix is invertible.

(2) So write

$$\left[\begin{array}{c|c} A & B \\ \hline B' & 0 \end{array}\right]^{-1} = \left[\begin{array}{c|c} C & D \\ \hline D' & E \end{array}\right]$$

and observe that

$$\left[\begin{array}{c|c} C & D \\ \hline D' & E \end{array}\right]\left[\begin{array}{c|c} A & B \\ \hline B' & 0 \end{array}\right] = \left[\begin{array}{c|c} I_n & 0 \\ \hline 0 & I_m \end{array}\right],$$

and that C and E are symmetric. (The inverse of a symmetric matrix is symmetric.) Expanding this yields

$$CA + DB' = I \tag{10}$$
$$CB = 0 \tag{11}$$
$$D'A + EB' = 0 \tag{12}$$
$$D'B = I \tag{13}$$

Now postmultiply (10) by C to get

$$CAC + D \underbrace{B'C}_{\substack{=0 \\ \text{by (11)}}} = C,$$

so

$$CAC = C. \tag{14}$$

By (11), for any x, we have $B'Cx = 0$, so Cx is orthogonal to each column of B. That is, Cx satisfies the constraints, so $(x'C)A(Cx) \leqslant 0$ with < 0 if $Cx \neq 0$. So by (14), $x'Cx = x'CACx \leqslant 0$ with < 0 if $Cx \neq 0$. That is, C is negative semidefinite.

To see that C has rank $n-m$, we show that $Cx = 0$ if and only if x is a linear combination of the columns of the m independent columns of B. Equation (11) already implies that $x = Bz$ implies $Cx = 0$. Now suppose $Cx = 0$. Premultiply (10) by x' to get

$$x'CA + x'DB' = x'.$$

Thus $x'C = 0$ implies $(x'D)B' = x'$, or $x = Bz$, where $z = Dx$.

Thus $Cx = 0$ if and only if x is a linear combination of the columns of B. Therefore the null space of C has dimension equal to the rank of B, which is m, so the rank of C equals $n - m$. ∎

The next result is a partial converse to Theorem 8.

Theorem 9 *Suppose A is an $n \times n$ symmetric matrix that is negative semidefinite under orthogonality constraints for the linearly independent vectors $b^1, \ldots, b^m$. That is, $x'Ax \leqslant 0$ for all nonzero x satisfying $B'x = 0$, where B is the $n \times m$ matrix whose j^{th} column is b^j. Suppose also that the matrix*

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

is invertible. Then A is actually negative definite under constraint. That is, $x'Ax < 0$ for all nonzero x satisfying $B'x = 0$.

Note that if B has full rank, then there are no nonzero x with $B'x = 0$. In that case the theorem is trivially true.

Proof: Suppose

$$\bar{x}'A\bar{x} = 0 \quad \text{and} \quad B'\bar{x} = 0.$$

Then $\bar{x}$ maximizes the quadratic form $\frac{1}{2}x \cdot Ax$ subject to the orthogonality constraints $B'x = 0$. Since the columns of B are independent, the constraint qualification is satisfied, so by the Lagrange Multiplier Theorem, there is a vector $\lambda \in \boldsymbol{R}^{\mathrm{m}}$ satisfying the first order conditions:

$$A\bar{x} + B\lambda = 0.$$

Thus

$$\left[\begin{array}{c|c} A & B \\ \hline B' & 0 \end{array}\right] \left[\begin{array}{c} \bar{x} \\ \hline \lambda \end{array}\right] = \left[\begin{array}{c} A\bar{x} + B\lambda \\ \hline B'\bar{x} \end{array}\right] = \left[\begin{array}{c} 0 \\ \hline 0 \end{array}\right].$$

Since $\left[\begin{array}{c|c} A & B \\ \hline B' & 0 \end{array}\right]$ is invertible, we see that $\bar{x} = 0$ (and $\lambda = 0$). Thus $B'x = 0$ and $x \neq 0$ imply $x'Ax < 0$. ∎

2.2 Determinantal conditions

Now consider the problem of maximizing the quadratic form $Q(x) = \frac{1}{2}x'Ax$ over the unit sphere subject to the orthogonality constraints $b^1 \cdot x = 0, \dots, b^m \cdot x = 0$. As in the proof of Proposition 6, we conclude that if x^* is a constrained maximizer, then there exist Lagrange multipliers λ^*, $\mu_1^*, \dots, \mu_m^*$ satisfying the first-order conditions

$$Ax^* - \lambda^* x + \mu_1^* b^1 + \cdots + \mu_m^* b^m = 0. \tag{15}$$

(Here we wrote the unit sphere constraint as $\frac{1}{2}(1 - x \cdot x) = 0$ to avoid unsightly fractions.) Premultiplying equation (15) by x^*, and using the fact that x^* is orthogonal to each b^i, we get

$$Q(x^*) = x^* \cdot Ax^* = \lambda^* x^* \cdot x^* = \lambda^*.$$

That is, the Lagrange multiplier λ^* is the maximum value of Q.

We can combine equation (15) with the orthogonality conditions in one big matrix equation:

$$\left[\begin{array}{c|c} A - \lambda^* I & B \\ \hline B' & 0 \end{array}\right] \left[\begin{array}{c} x^* \\ \hline \mu^* \end{array}\right] = \left[\begin{array}{c} 0 \\ \hline 0 \end{array}\right],$$

where B is the matrix whose columns are $b^1, \dots, b^m$ and μ^* is the vector with components $\mu_1^*, \dots, \mu_m^*$. Since x^* is nonzero (it lies on the unit sphere), the matrix $\left[\begin{array}{c|c} A - \lambda^* I & B \\ \hline B' & 0 \end{array}\right]$ must be singular, so

$$\det \left[\begin{array}{c|c} A - \lambda^* I & B \\ \hline B' & 0 \end{array}\right] = 0.$$

The next result is due to Hancock [9, pp. 105–114], who attributes the approach to Lagrange.

Proposition 10 (Hancock) *Let A be an $n \times n$ symmetric matrix and let $\{b^1, \dots, b^m\}$ be linearly independent. Let*

$$f(\lambda) = \det \left[\begin{array}{c|c} A - \lambda I & B \\ \hline B' & 0 \end{array}\right].$$

If all the coefficients of f have the same sign, then A is negative semidefinite under constraint.

If the coefficients of f alternate in sign, then A is positive semidefinite under constraint. (Here we must consider the zero coefficients to be alternating in sign.)

If in addition, $f(0) = \det \left[\begin{array}{c|c} A & B \\ \hline B' & 0 \end{array}\right] \neq 0$, then A is actually definite under constraint.

Proof: Even without resort to Descartes' infamous Rule of Signs the following fact is easy to see: *If all the nonzero coefficients of a nonzero polynomial f (of degree at least one) have the same sign, then f has no strictly positive roots.* For if all the coefficients of a polynomial f are nonnegative, then $f(0) \geqslant 0$ and f is strictly increasing on $(0, \infty)$, so it has no positive roots. Likewise if all the coefficients are nonpositive, then $f(0) \leqslant 0$ and f is strictly decreasing on $(0, \infty)$, so it has no positive roots. Trivially, if $f(0) \neq 0$, then 0 is not a root.

From the discussion preceding the proposition, λ^*, the maximum value of $x'Ax$ on the unit sphere, is a root of f. If the coefficients of f do not change sign, then $\lambda^* \leqslant 0$. That is, A is negative semidefinite under constraint, and is actually definite if $f(0) \neq 0$.

The results on positive (semi)definiteness follow from the fact that λ^* is a negative root of $f(\lambda)$ if and only if $-\lambda^*$ is a positive root of $f(-\lambda)$. ∎

The problem with applying Hancock's result is that he does not provide a simple formula for the coefficients.

2.3 Bordered matrices and quadratic forms

We define the matrices of the form

$$\begin{bmatrix} a_{11} & \dots & a_{1r} & b_1^1 & \dots & b_1^m \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{r,1} & \dots & a_{rr} & b_r^1 & \dots & b_r^m \\ b_1^1 & \dots & b_r^1 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ b_1^m & \dots & b_r^m & 0 & \dots & 0 \end{bmatrix}$$

to be $\boldsymbol{r}^{\mathbf{th}}$ **order bordered minors of $\boldsymbol{A}$**. Note that the r refers to the number of rows and columns from A. The actual r^{th} order minor has $m + r$ rows and columns, where m is the number of constraint vectors. The proof of the following result may be found in Debreu [5, Theorems 4 and 5] or Mann [10]. Note that Mann errs in the statement of part 2. A proof may also be found sketched in Samuelson [14, pp. 376–378].

Theorem 11 *Let A be an $n \times n$ symmetric matrix and let $\{b^1, \dots, b^m\}$ be linearly independent.*

1. *A is positive definite under the orthogonality constraints $b^1, \dots, b^m$ if and only if*

$$(-1)^m \begin{vmatrix} a_{11} & \dots & a_{1r} & b_1^1 & \dots & b_1^m \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{r,1} & \dots & a_{rr} & b_r^1 & \dots & b_r^m \\ b_1^1 & \dots & b_r^1 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ b_1^m & \dots & b_r^m & 0 & \dots & 0 \end{vmatrix} > 0$$

   for $r = m+1, \dots, n$. That is, if and only if every r^{th}-order NW bordered principal minor has sign $(-1)^m$ for $r > m$.

2. *A is negative definite under the orthogonality constraints $b^1, \dots, b^m$ if*

and only if

$$(-1)^r \begin{vmatrix} a_{11} & \dots & a_{1r} & b_1^1 & \dots & b_1^m \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{r,1} & \dots & a_{rr} & b_r^1 & \dots & b_r^m \\ b_1^1 & \dots & b_r^1 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ b_1^m & \dots & b_r^m & 0 & \dots & 0 \end{vmatrix} > 0$$

for $r = m+1, \dots, n$. *That is, if and only if every* r^{th}*-order NW bordered principal minor has sign* $(-1)^r$ *for* $r > m$.

Note that for positive definiteness under constraint all the NW bordered principal minors of order greater than m have the same sign, the sign depending on the number of constraints. For negative definiteness the NW bordered principal minors alternate in sign. For the case of one constraint ($m = 1$) if A is positive definite under constraint, then these minors are negative. Again with one constraint, if A is negative definite under constraint, then the minors of even order are positive and of odd order are negative.

To see how to derive statement (2) from statement (1), observe that A is negative definite under constraint if and only if $-A$ is positive definite under constraint, which by statement (1) is equivalent to

$$(-1)^m \begin{vmatrix} -a_{11} & \dots & -a_{1r} & b_1^1 & \dots & b_1^m \\ \vdots & & \vdots & \vdots & & \vdots \\ -a_{r,1} & \dots & -a_{rr} & b_r^1 & \dots & b_r^m \\ b_1^1 & \dots & b_r^1 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ b_1^m & \dots & b_r^m & 0 & \dots & 0 \end{vmatrix} > 0$$

for $r = m + 1, \dots, n$. But multiplying the first r rows by -1 multiplies the determinant by $(-1)^r$, so

$$(-1)^m \begin{vmatrix} -a_{11} & \dots & -a_{1r} & b_1^1 & \dots & b_1^m \\ \vdots & & \vdots & \vdots & & \vdots \\ -a_{r,1} & \dots & -a_{rr} & b_r^1 & \dots & b_r^m \\ b_1^1 & \dots & b_r^1 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ b_1^m & \dots & b_r^m & 0 & \dots & 0 \end{vmatrix}$$

$$= (-1)^{m+r} \begin{vmatrix} a_{11} & \dots & a_{1r} & -b_1^1 & \dots & -b_1^m \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{r,1} & \dots & a_{rr} & -b_r^1 & \dots & -b_r^m \\ b_1^1 & \dots & b_r^1 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ b_1^m & \dots & b_r^m & 0 & \dots & 0 \end{vmatrix}$$

and then multiplying the last m columns by -1 multiplies the determinant by $(-1)^m$,

$$= (-1)^{2m+r} \begin{vmatrix} a_{11} & \dots & a_{1r} & b_1^1 & \dots & b_1^m \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{r,1} & \dots & a_{rr} & b_r^1 & \dots & b_r^m \\ b_1^1 & \dots & b_r^1 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ b_1^m & \dots & b_r^m & 0 & \dots & 0 \end{vmatrix}$$

Since $(-1)^{2m+r} = (-1)^r$, statement (2) follows.

References

[1] T. W. Anderson. 1958. *An introduction to multivariate statistical analysis*. A Wiley Publication in Mathematical Statistics. New York: Wiley.

[2] T. M. Apostol. 1969. *Calculus*, 2d. ed., volume 2. Waltham, Massachusetts: Blaisdell.

[3] C. Carathéodory. 1982. *Calculus of variations*, 2d. ed. New York: Chelsea. This was originally published in 1935 in two volumes by B. G. Teubner in Berlin as *Variationsrechnung und Partielle Differentialgleichungen erster Ordnung*. In 1956 the first volume was edited and updated by E. Hölder. The revised work was translated by Robert B. Dean and Julius J. Brandstatter and published in two volumes as *Calculus of variations and partial differential equations of the first order* by Holden-Day in 1965–66. The Chelsea second edition combines and revises the 1967 edition.

[4] Y. Chabrillac and J. P. Crouzeix. 1984. Definiteness and semidefiniteness of quadratic forms revisited. *Linear Algebra and its Applications* 63:283–292. DOI: 10.1016/0024-3795(84)90150-2

[5] G. Debreu. 1952. Definite and semidefinite quadratic forms. *Econometrica* 20(2):295–300. http://www.jstor.org/stable/1907852

[6] W. E. Diewert and A. D. Woodland. 1977. Frank Knight's theorem in linear programming revisited. *Econometrica* 45(2):375–398. http://www.jstor.org/stable/1911216

[7] J. Franklin. 2000. *Matrix theory*. Mineola, New York: Dover. Reprint of 1968 edition published by Prentice-Hall, Englewood Cliffs, N.J.

[8] F. R. Gantmacher. 1959. *Matrix theory*, volume 1. New York: Chelsea.

[9] H. Hancock. 1960. *Theory of maxima and minima*. New York: Dover. Reprint of 1917 edition published by Ginn–Blaisdell.

[10] H. B. Mann. 1943. Quadratic forms with linear constraints. *American Mathematical Monthly* 50(7):430–433. Reprinted in [11]. http://www.jstor.org/stable/2303666

[11] P. Newman, ed. 1968. *Readings in mathematical economics I: Value theory*. Baltimore: Johns Hopkins Press.

[12] J. P. Quirk. 1976. *Intermediate microeconomics: Mathematical notes*. Chicago: Science Research Associates.

[13] C. R. Rao. 1973. *Linear statistical inference and its applications*, 2d. ed. Wiley Series in Probability and Mathematical Statistics. New York: Wiley.

[14] P. A. Samuelson. 1965. *Foundations of economic analysis*. New York: Athenaeum. Reprint of the 1947 edition published by Harvard University Press.