ECE 602 Fall 2010

Quadratic Forms and the Kronecker Product

by

Stanislaw H. Żak

September 28, 2010

Quadratic Forms

A real *quadratic form* is the product $\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x}$, where $\boldsymbol{Q}$ is a specified $n \times n$ matrix with real coefficients and $\boldsymbol{x}$ is an $n \times 1$ real, column, vector. Without loss of generality, the matrix $\boldsymbol{Q}$ can be taken to be symmetric, that is, $\boldsymbol{Q} = \boldsymbol{Q}^\top$. Indeed, if $\boldsymbol{Q}$ is not symmetric, we can always replace it with a symmetric matrix without changing the quadratic form values. Indeed, because $z = \boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x}$ is a scalar, hence $z^\top = z$, and we have

$$\left(\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x}\right)^\top = \boldsymbol{x}^\top \boldsymbol{Q}^\top \boldsymbol{x} = \boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x}.$$

Therefore, we can write

$$\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} = \boldsymbol{x}^\top \left(\frac{1}{2}\boldsymbol{Q} + \frac{1}{2}\boldsymbol{Q}^\top\right)\boldsymbol{x} = \boldsymbol{x}^\top \left(\frac{\boldsymbol{Q}+\boldsymbol{Q}^\top}{2}\right)\boldsymbol{x},$$

where $\frac{1}{2}\left(\boldsymbol{Q}+\boldsymbol{Q}^\top\right)$ is a symmetric matrix. Thus, the quadratic form values are unchanged if $\boldsymbol{Q}$ is replaced with the symmetric matrix $\frac{1}{2}\left(\boldsymbol{Q}+\boldsymbol{Q}^\top\right)$. From now on, in our study of quadratic forms, we assume that $\boldsymbol{Q}$ is symmetric.

A quadratic form $\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x}$, or equivalently, the matrix $\boldsymbol{Q}$, is called *positive semi-definite* if $\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} \geq 0$ for all $\boldsymbol{x}$. It is *positive definite* if $\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} > 0$ for all $\boldsymbol{x} \neq \boldsymbol{0}$. If $\boldsymbol{Q}$ is positive semi-definite, then we write $\boldsymbol{Q} \geq 0$, if it is positive definite, we write $\boldsymbol{Q} > 0$. We write $\boldsymbol{Q}_1 \geq \boldsymbol{Q}_2$ to mean that $\boldsymbol{Q}_1 - \boldsymbol{Q}_2 \geq 0$. We say that $\boldsymbol{Q}$ is *negative semi-definite* if $\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} \leq 0$ for all $\boldsymbol{x}$, it is *negative definite* if $\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} < 0$ for all $\boldsymbol{x} \neq \boldsymbol{0}$. Note that $\boldsymbol{Q} \leq 0$ if and only if $-\boldsymbol{Q} \geq 0$.

We now present tests for definiteness properties of symmetric matrices. We state these tests in terms of *principal minors* of a given real, symmetric, $n \times n$ matrix $\boldsymbol{Q}$ with elements

q_{ij}. We will also give tests for definiteness in terms of eigenvalues of $\boldsymbol{Q}$. The principal minors of $\boldsymbol{Q}$ are $\det \boldsymbol{Q}$ itself and the determinants of submatrices of $\boldsymbol{Q}$ obtained by removing successively an i-th row and i-th column. Specifically, for $p = 1, 2, \ldots, n$, the principal minors are

$$\Delta_p \begin{pmatrix} i_1, & i_2, & \ldots, & i_p \\ i_1, & i_2, & \ldots, & i_p \end{pmatrix} = \det \begin{bmatrix} q_{i_1 i_1} & q_{i_1 i_2} & \cdots & q_{i_1 i_p} \\ q_{i_2 i_1} & q_{i_2 i_2} & \cdots & q_{i_2 i_p} \\ \vdots & \vdots & \ddots & \vdots \\ q_{i_p i_1} & q_{i_p i_2} & \cdots & q_{i_p i_p} \end{bmatrix} \quad 1 \le i_1 < i_2 < \cdots < i_p \le n.$$

The above are called the principal minors of order p.

In the following, we use the fact that if $\boldsymbol{Q} = \boldsymbol{Q}^\top$, then its eigenvalues are all real.

Theorem 1 *The following are equivalent:*

(i) The symmetric matrix $\boldsymbol{Q}$ is positive semi-definite.

(ii) Principal minors of $\boldsymbol{Q}$ are nonnegative, that is, for $p = 1, 2, \ldots, n$,

$$\Delta_p \begin{pmatrix} i_1, & i_2, & \ldots, & i_p \\ i_1, & i_2, & \ldots, & i_p \end{pmatrix} \ge 0, \quad 1 \le i_1 < i_2 < \cdots < i_p \le n.$$

(iii) Eigenvalues of $\boldsymbol{Q}$ are nonnegative,

$$\lambda_i(\boldsymbol{Q}) \ge 0, \quad i = 1, 2, \ldots, n.$$

We now give two tests for $\boldsymbol{Q} = \boldsymbol{Q}^\top$ to be positive definite; one in terms of its leading principal minors, and the other in terms of its eigenvalues. The *leading principal minors* of $\boldsymbol{Q}$ are

$$\Delta_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = q_{11}, \quad \Delta_2 \begin{pmatrix} 1, & 2 \\ 1, & 2 \end{pmatrix} = \det \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix},$$

$$\Delta_3 \begin{pmatrix} 1, & 2, & 3 \\ 1, & 2, & 3 \end{pmatrix} = \det \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix}, \ldots, \Delta_n \begin{pmatrix} 1, & 2, & \ldots, & n \\ 1, & 2, & \ldots, & n \end{pmatrix} = \det \boldsymbol{Q}.$$

Theorem 2 *The following are equivalent:*

(i) The symmetric matrix $\boldsymbol{Q}$ is positive definite.

(ii) Leading principal minors of $\boldsymbol{Q}$ are positive, that is,

$$\Delta_p \begin{pmatrix} 1, & 2, & \ldots, & p \\ 1, & 2, & \ldots, & p \end{pmatrix} > 0, \quad p = 1, 2, \ldots, n.$$

(iii) Eigenvalues of $\boldsymbol{Q}$ are positive,

$$\lambda_i(\boldsymbol{Q}) > 0, \quad i = 1, 2, \ldots, n.$$

If a given matrix $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ is symmetric, then its eigenvalues are all real and there exists a unitary matrix $\boldsymbol{U}$ such that

$$\boldsymbol{U}^\top \boldsymbol{Q} \boldsymbol{U} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \boldsymbol{\Lambda}. \tag{1}$$

Using the above, one can show that a symmetric positive semi-definite matrix $\boldsymbol{Q}$ has a positive semi-definite symmetric square root, denoted $\boldsymbol{Q}^{1/2}$ or $\sqrt{\boldsymbol{Q}}$, that satisfies

$$\boldsymbol{Q}^{1/2}\boldsymbol{Q}^{1/2} = \boldsymbol{Q}. \tag{2}$$

Indeed, it follows from Theorem 1 that if $\boldsymbol{Q} = \boldsymbol{Q}^\top \geq 0$ then its eigenvalues are all nonnegative. Hence,

$$\boldsymbol{\Lambda}^{1/2} = \begin{bmatrix} \lambda_1^{1/2} & 0 & \cdots & 0 \\ 0 & \lambda_2^{1/2} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^{1/2} \end{bmatrix} \tag{3}$$

is well defined. Using the fact that $\boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}_n$, we can define

$$\boldsymbol{Q}^{1/2} = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}^\top, \tag{4}$$

which has the desired properties.

If $\boldsymbol{Q} = \boldsymbol{Q}^\top$ is positive semi-definite, and not positive definite, then some of its eigenvalues are equal to zero. Suppose that $\boldsymbol{Q}$ has r nonzero eigenvalues. Then, there is a unitary matrix, say $\boldsymbol{U}$, such that

$$\boldsymbol{U}^\top \boldsymbol{Q} \boldsymbol{U} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{5}$$

Let $\boldsymbol{V} \in \mathbb{R}^{n \times r}$ be a matrix that consists of the first r columns of the matrix $\boldsymbol{U}$, and let

$$\boldsymbol{C} = \begin{bmatrix} \lambda_1^{1/2} & 0 & \cdots & 0 \\ 0 & \lambda_2^{1/2} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_r^{1/2} \end{bmatrix} \boldsymbol{V}^\top. \tag{6}$$

Note that the matrix $\boldsymbol{C}$ is of full rank. We have

$$\boldsymbol{Q} = \boldsymbol{C}^\top \boldsymbol{C}. \tag{7}$$

We refer to (7) as a full rank factorization of $\boldsymbol{Q}$ because $\boldsymbol{C}$ is of full rank.

The Kronecker Product

Let $\boldsymbol{A}$ be an $m \times n$ and $\boldsymbol{B}$ be a $p \times q$ matrices. The *Kronecker product* of $\boldsymbol{A}$ and $\boldsymbol{B}$, denoted $\boldsymbol{A} \otimes \boldsymbol{B}$, is an $mp \times nq$ matrix defined by

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \cdots & a_{2n}\boldsymbol{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\boldsymbol{B} & a_{m2}\boldsymbol{B} & \cdots & a_{mn}\boldsymbol{B} \end{bmatrix}. \tag{8}$$

Thus, the matrix $\boldsymbol{A} \otimes \boldsymbol{B}$ consists of mn blocks. Using the definition of the Kronecker product, we can verify the following two properties:

$$(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = \boldsymbol{A}\boldsymbol{C} \otimes \boldsymbol{B}\boldsymbol{D}, \tag{9}$$

$$(\boldsymbol{A} \otimes \boldsymbol{B})^\top = \boldsymbol{A}^\top \otimes \boldsymbol{B}^\top. \tag{10}$$

Let

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}_1, & \boldsymbol{x}_2, & \ldots, & \boldsymbol{x}_m \end{bmatrix}$$

be a $n \times m$ matrix, where $\boldsymbol{x}_i$, $i = 1, 2, \ldots, m$ are the columns of $\boldsymbol{X}$. Each $\boldsymbol{x}_i$ consists of n elements. Then, the *vectorization operator* or *stacking operator* is defined as

$$\begin{aligned} \operatorname{vec}(\boldsymbol{X}) &= \begin{bmatrix} \boldsymbol{x}_1^\top, & \boldsymbol{x}_2^\top, & \ldots, & \boldsymbol{x}_m^\top \end{bmatrix}^\top \\ &= \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{n1} & x_{12} & x_{22} & \cdots & x_{n2} & \cdots & x_{1m} & x_{2m} & \cdots & x_{nm} \end{bmatrix}^\top \end{aligned} \tag{11}$$

is the column nm-vector formed from the columns of $\boldsymbol{X}$ taken in order. Let now $\boldsymbol{A}$ be an $n \times n$, and $\boldsymbol{C}$ and $\boldsymbol{X}$ be $n \times m$ matrices. Then, the matrix equation

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{C} \tag{12}$$

can be written as

$$(\boldsymbol{I}_m \otimes \boldsymbol{A}) \operatorname{vec}(\boldsymbol{X}) = \operatorname{vec}(\boldsymbol{C}). \tag{13}$$

Let $\boldsymbol{B}$ be an $m \times m$ matrix, then the matrix equation

$$\boldsymbol{X}\boldsymbol{B} = \boldsymbol{C} \tag{14}$$

can be written as

$$\left(\boldsymbol{B}^\top \otimes \boldsymbol{I}_n\right) \operatorname{vec}(\boldsymbol{X}) = \operatorname{vec}(\boldsymbol{C}). \tag{15}$$

Using the above two facts, we can verify that the Sylvester matrix equation,

$$\boldsymbol{A}\boldsymbol{X} + \boldsymbol{X}\boldsymbol{B} = \boldsymbol{C}, \tag{16}$$

where $\boldsymbol{A}$ is $n \times n$, $\boldsymbol{B}$ is $m \times m$, and $\boldsymbol{C}$ is $n \times m$, can be written as

$$\left(\boldsymbol{I}_m \otimes \boldsymbol{A} + \boldsymbol{B}^\top \otimes \boldsymbol{I}_n\right) \operatorname{vec}(\boldsymbol{X}) = \operatorname{vec}(\boldsymbol{C}). \tag{17}$$

Let λ_i, $\boldsymbol{v}_i$ be the eigenvalues and eigenvectors, respectively of $\boldsymbol{A}$, and μ_j and $\boldsymbol{w}_j$ the eigenvalues and eigenvectors of $m \times m$ matrix $\boldsymbol{B}$. Then,

$$\begin{aligned} (\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{v}_i \otimes \boldsymbol{w}_j) &= \boldsymbol{A}\boldsymbol{v}_i \otimes \boldsymbol{B}\boldsymbol{w}_j \\ &= \lambda_i \boldsymbol{v}_i \otimes \mu_j \boldsymbol{w}_j \\ &= \lambda_i \mu_j (\boldsymbol{v}_i \otimes \boldsymbol{w}_j). \end{aligned} \tag{18}$$

Thus, the eigenvalues of $\boldsymbol{A} \otimes \boldsymbol{B}$ are $\lambda_i \mu_j$, and their respective eigenvectors are $\boldsymbol{v}_i \otimes \boldsymbol{w}_j$ for $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$.

We use the above result to study the matrix equation given by (16). We represent (16) as

$$\boldsymbol{M} \operatorname{vec}(\boldsymbol{X}) = \operatorname{vec}(\boldsymbol{C}), \tag{19}$$

where

$$\boldsymbol{M} = \boldsymbol{I}_m \otimes \boldsymbol{A} + \boldsymbol{B}^\top \otimes \boldsymbol{I}_n. \tag{20}$$

The solution to (19) is unique if, and only if, the $mn \times mn$ matrix $\boldsymbol{M}$ is nonsingular. To find the condition for this to hold, consider the following matrix

$$\left(\boldsymbol{I}_m + \varepsilon \boldsymbol{B}^\top\right) \otimes (\boldsymbol{I}_n + \varepsilon \boldsymbol{A}) = \boldsymbol{I}_m \otimes \boldsymbol{I}_n + \varepsilon \boldsymbol{M} + \varepsilon^2 \boldsymbol{B}^\top \otimes \boldsymbol{A} \tag{21}$$

whose eigenvalues are

$$(1 + \varepsilon \mu_j)(1 + \varepsilon \lambda_i) = 1 + \varepsilon (\mu_j + \lambda_i) + \varepsilon^2 \mu_j \lambda_i \tag{22}$$

because for a square matrix $\boldsymbol{Q}$,

$$\lambda_i (\boldsymbol{I}_n + \varepsilon \boldsymbol{Q}) = 1 + \varepsilon \lambda_i (\boldsymbol{Q}).$$

Comparing terms in ε in (21) and (22), we conclude that the eigenvalues of $\boldsymbol{M}$ are $\lambda_i + \mu_j$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$. Hence, $\boldsymbol{M}$ is nonsingular if and only if

$$\lambda_i + \mu_j \neq 0. \tag{23}$$

The above is the necessary and sufficient condition for the solution $\boldsymbol{X}$ of the matrix equation $\boldsymbol{AX} + \boldsymbol{XB} = \boldsymbol{C}$ to be unique.

Another useful identity involving the Kronecker product is

$$\operatorname{vec}(\boldsymbol{ABC}) = \left(\boldsymbol{C}^\top \otimes \boldsymbol{A}\right) \operatorname{vec}(\boldsymbol{B}) \tag{24}$$

For further information on the subject of the Kronecker product, we refer the reader to Brewer [1].

References

[1] J. W. Brewer. Kronecker products and matrix calculus in system theory. *IEEE Transactions on Circuits and Systems*, CAS-25(9):772–781, September 1978.