mention is made of Carathéodory himself. There are a few calculus texts, including Boyce and DiPrima [1], which present the variation mentioned at the end of the previous section.

(Note: Young [13] contains interesting biographical information on Carathéodory, from the viewpoint of one of his students and strong admirers.)

Acknowledgement. Thanks to the referee for several helpful suggestions.

REFERENCES

1. William Boyce and Richard DiPrima, Calculus, Wiley, New York, 1988.
2. Constantin Carathéodory, Theory of Functions of a Complex Variable, vol. 1, Chelsea, New York, 1954.
3. ______, Vorlesungen über Reele Functionen, Chelsea Publishing Company, New York, 1968.
4. C. H. Edwards, Jr., The Historical Development of the Calculus, Springer-Verlag, New York, 1979.
5. Robert Ellis and Denny Gulick, Calculus with Analytic Geometry, Harcourt Brace Jovanovich, New York, 1978.
6. Philip Gillett, Calculus and Analytic Geometry, third ed., D. C. Heath, Lexington, Mass., 1988.
7. Judith V. Grabiner, The Origin of Cauchy's Rigorous Calculus, MIT Press, Cambridge, Mass., 1981.
8. Maurice Heins, Complex Function Theory, Academic Press, New York, 1968.
9. Louis Leithold, The Calculus with Analytic Geometry, fifth ed., Harper and Row, New York, 1986.
10. M. H. Protter and C. B. Morrey, A First Course in Real Analysis, Springer-Verlag, New York, 1977.
11. Edwin Purcell and Dale Varburg, Calculus with Analytic Geometry, fifth ed., Prentice-Hall, Englewood Cliffs, N.J., 1987.
12. Maxwell Rosenlicht, Introduction to Analysis, Scott Foresman, Glenview, Ill., 1968.
13. Laurence Young, Mathematicians and Their Times, North-Holland, Amsterdam, 1981.

Positive Definite Matrices and Sylvester's Criterion

George T. Gilbert
Department of Mathematics, Texas Christian University,
Fort Worth, TX 76129

Sylvester's criterion states that a symmetric (more generally, Hermitian) matrix is positive definite if and only if its principal minors are all positive. The theorem is especially useful when employing the Second Derivative Test for local extrema of functions of several variables. When the proof of the sufficiency of positive principal minors is included in a linear algebra text, it is usually based on Gaussian elimination [**4**, pp. 331–332] or on the reduction theory for quadratic forms [**2**, pp. 328–329]. In either case, the heart of such proofs is a matrix computation. This note presents a proof which, while still very dependent on matrices, makes more use of ideas from the theory of vector spaces. In a one semester course in linear algebra, one often has the feeling of developing the machinery of abstract vector spaces without ever really making use of it. The proof here is short, requires only concepts that are standard in beginning linear algebra, and provides an opportunity for the student to see the abstract theory applied to a concrete problem. A less direct proof in a similar spirit is given in [**3**].

We begin by stating the prerequisite theorems and definitions. The first theorem we shall need is the following:

THEOREM 1. *Let W_1 and W_2 be finite-dimensional subspaces of a vector space V. Then*

$$\dim(W_1 + W_2) + \dim(W_1 \cap W_2) = \dim W_1 + \dim W_2.$$

Proof. One simply extends a basis for $W_1 \cap W_2$ to a basis for W_1 and to a basis for W_2. The bases combine to form a basis for $W_1 + W_2$. See [**2**, p. 46] or [**4**, pp. 199–200] for details.

Since the determinants of similar matrices are equal, we also have

PROPOSITION. *The determinant of a diagonalizable matrix is the product of its eigenvalues.*

The next theorem provides one of many opportunities for nontrivial inductive proofs in a linear algebra course.

THEOREM 2 (Spectral Theorem). *Let A be an $n \times n$ real, symmetric matrix. Then all eigenvalues of A are real and there exists an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A.*

Proof. See [**2**, pp. 312–314] or [**4**, pp. 295–296] for a proof.

Having stated the background we require, we proceed to positive definite matrices.

DEFINITION. A real, symmetric matrix is *positive definite* if for every nonzero vector in $\mathbb{R}^n$, $\mathbf{v}^t A \mathbf{v} > 0$.

The proof of the next theorem is a routine exercise.

THEOREM 3. *A real, symmetric matrix A is positive definite if and only if all its eigenvalues are positive.*

Once we define principal minors, we turn to the proof of Sylvester's criterion.

DEFINITION. Let A be an $n \times n$ matrix. For $1 \leq k \leq n$, the *kth principal submatrix* of A is the $k \times k$ submatrix formed from the first k rows and first k columns of A. Its determinant is the *kth principal minor*.

THEOREM (Sylvester's Criterion). *A real, symmetric matrix is positive definite if and only if all its principal minors are positive.*

It is easy to prove that positive principal minors are necessary by showing that each kth principal submatrix is positive definite, hence has positive determinant by Theorem 3 and the Proposition. Our proof that positive principal minors implies a matrix is positive definite proceeds through two lemmas, instructive in their own rights. At each stage we use vector space methods, although, of course, parallel matrix proofs are possible.

Lemma 1. *Let* $\mathbf{v}_1, \dots, \mathbf{v}_n$ *be a basis of a vector space* V. *Suppose* W *is a* k*-dimensional subspace of* V. *If* $m < k$, *there exists a nonzero vector in* W *which is a linear combination of* $\mathbf{v}_{m+1}, \dots, \mathbf{v}_n$.

Proof. Since $\dim W + (n - m) = k + n - m > n$, W has nontrivial intersection with the span of $\mathbf{v}_{m+1}, \dots, \mathbf{v}_n$ by Theorem 1.

The difference between our proof and others lies mainly in the next lemma, which is nearly a special case of the Courant-Fischer "min-max theorem" [**3**, p. 179].

Lemma 2. *Let* A *be an* $n \times n$ *real, symmetric matric. If* $\mathbf{w}^t A \mathbf{w} > 0$ *for all nonzero vectors* $\mathbf{w}$ *in a* k*-dimensional subspace* W *of* $\mathbb{R}^n$, *then* A *has at least* k *positive eigenvalues (counting multiplicity).*

Proof. Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors for A. Let $\lambda_1, \dots, \lambda_n$ be the corresponding eigenvalues. Suppose further that the first m eigenvalues are positive and the rest are not. If $m < k$, then Lemma 1 implies there exists a nonzero vector $\mathbf{w}$ in W which may be written

$$\mathbf{w} = c_{m+1}\mathbf{v}_{m+1} + \cdots + c_n \mathbf{v}_n.$$

We have

$$\mathbf{w}^t A \mathbf{w} = \sum_{j,k=m+1}^{n} c_j c_k \lambda_k \mathbf{v}_j^t \cdot \mathbf{v}_k = c_{m+1}^2 \lambda_{m+1} + \cdots + c_n^2 \lambda_n,$$

since $\mathbf{v}_1, \dots, \mathbf{v}_n$ is an orthonormal set. But then $\mathbf{w}^t A \mathbf{w} \le 0$, a contradiction. Thus $m \ge k$, as desired.

We complete the proof of sufficiency in Sylvester's criterion by induction. For $n = 1$, the result is trivial. Assume the sufficiency of positive principal minors for $(n-1) \times (n-1)$ real, symmetric matrices. If A is an $n \times n$ real, symmetric matrix with positive principal minors, then, by the inductive hypothesis, its $(n-1)$st principal submatrix is positive definite. Let W be the $(n-1)$-dimensional subspace of $\mathbb{R}^n$ consisting of those vectors whose last coordinate is zero. Then for any nonzero vector $\mathbf{w}$ in W, $\mathbf{w}^t A \mathbf{w} > 0$. Lemma 2 now implies that A has at least $n - 1$ positive eigenvalues (counting multiplicity). We apply the Proposition above and the fact that $\det A > 0$ to conclude that A has n positive eigenvalues. This completes the proof of Sylvester's criterion.

Acknowledgements. The author wishes to thank the referee for suggesting several improvements and for bringing [**3**] to his attention.

REFERENCES

1. N. V. Efimov, E. R. Rozendorn, Linear Algebra and Multidimensional Geometry, Mir Publishers, Moscow, 1975.
2. K. Hoffman and R. Kunze, Linear Algebra, 2nd ed., Prentice-Hall, Englewood Cliffs, N.J., 1971.
3. R. A. Horn and C. R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.
4. G. Strang, Linear Algebra and Its Applications, 3rd ed., Harcourt Brace Jovanovich, Inc., San Diego, 1988.