JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 41, 164–169 (1973)

On the Ostrowski–Schneider Inertia Theorem

H. K. WIMMER

Technische Hochschule Graz, Austria

Submitted by Ky Fan

1.

In [1] the inertia, $\mathrm{In}(A)$, of a matrix A is defined as a triple of nonnegative integers (π, ν, δ). $\pi = \pi(A)$ is the number of eigenvalues of A with positive real parts, there are $\nu = \nu(A)$ eigenvalues of A with negative real parts and $\delta = \delta(A)$ is the number of eigenvalues of A on the imaginary axis. $M > 0$ ($M < 0$) denotes a positive (negative) definite Hermitian matrix M.

The first result relating the inertia of a matrix with a matrix inequality is due to Ljapunov.

THEOREM 1 [3]. *If all eigenvalues of A have positive real parts and H is a Hermitian matrix with*

$$A^*H + HA > 0, \tag{1}$$

then H is positive definite.

THEOREM 2 [3]. *If all eigenvalues of A have modulus less than 1 and H is a Hermitian matrix with*

$$H - A^*HA > 0,$$

then H is positive definite.

A generalization of Theorem 1 is the inertia theorem of Ostrowski and Schneider.

THEOREM 3 [1]. *If H is a Hermitian solution of*

$$A^*H + HA = C, \qquad C > 0, \tag{2}$$

then

$$\mathrm{In}(A) = \mathrm{In}(H).$$

This theorem was proved by Taussky [2] for $C = I$ and by Ostrowski and Schneider for general C. In this note we give a shorter proof of Theorem 3 by demonstrating that it is equivalent to

THEOREM 4. *Let*

$$G := \begin{pmatrix} G_{11} & G_{12} \\ G_{12}^* & G_{22} \end{pmatrix} \tag{3}$$

be a Hermitian $m \times m$ matrix. If the $p \times p$ matrix G_{11} is positive definite and the $n \times n$ matrix G_{22} is negative definite, $n + p = m$, then

$$\mathrm{In}(G) = (p, n, 0).$$

For the equation $H - A^*HA = C$, $C > 0$ an analogon to Theorem 3 will be proved. All proofs are given in Section 2.

Let $\{\lambda_1, ..., \lambda_k\}$ be the set of eigenvalues of the $n \times n$ matrix A. Then A can be written in the form [4]

$$A = \sum_{i=1}^{k} (\lambda_i I + N_i) P_i .$$

$\{N_i\}$ is a set of nilpotent matrices, $\{P_i\}$ is a set of projection matrices, such that

$$\sum_{i=1}^{k} P_i = I, \qquad P_i P_j = \delta_{ij} P_i , \qquad P_i N_j = N_j P_i = \delta_{ij} N_i .$$

We put

$$P_+ := \sum_{\mathrm{Re}\lambda_i > 0} P_i \qquad \text{and} \qquad P_- := \sum_{\mathrm{Re}\lambda_i < 0} P_i .$$

P_+ and P_- define projections of $\mathbf{C}^n$ on subspaces $P_+\mathbf{C}^n$ and $P_-\mathbf{C}^n$ of $\mathbf{C}^n$.

For a Hermitian solution of (1) the following theorem holds:

THEOREM 5. *If a Hermitian matrix H satisfies $A^*H + HA > 0$, then H is positive definite on $P_+\mathbf{C}^n$ and negative definite on $P_-\mathbf{C}^n$.*

In Theorem 9 we characterize matrices A which have the property that $A^*H + HA > 0$ for all H which are positive definite on $P_+\mathbf{C}^n$ and negative definite on $P_-\mathbf{C}^n$.

2.

Proof of Theorem 4. We put

$$T := \begin{pmatrix} I & -G_{11}^{-1}G_{12} \\ 0 & I \end{pmatrix}.$$

Then

$$T^*GT = \begin{pmatrix} G_{11} & 0 \\ 0 & -G_{12}^*G_{11}^{-1}G_{12} + G_{22} \end{pmatrix}.$$

Because of $G_{22} < 0$ and $G_{11} > 0$ we have

$$-G_{12}^*G_{11}^{-1}G_{12} + G_{22} < 0 \qquad \text{and} \qquad \operatorname{In}(T^*GT) = \operatorname{In}(G) = (p, n, 0).$$

THEOREM 6. *Theorem* 3 *and Theorem* 4 *are equivalent.*

Proof. (a) Theorem 4 $\Rightarrow$ Theorem 3. We assume (2). Let S be a matrix which transforms A to Jordan form.

$$B := S^{-1}AS = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}. \tag{4}$$

In B_1 are the blocks corresponding to eigenvalues with positive real parts, in B_2 the blocks corresponding to eigenvalues with negative real parts. Because of (2) there are no purely imaginary eigenvalues of A. If we define $Q := S^*CS$ and $M := S^*HS$, then S and S^* transform (2) into

$$B^*M + MB = Q, \qquad Q > 0. \tag{5}$$

M and Q are partitioned according to (4):

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^* & M_{22} \end{pmatrix}, \qquad Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{12}^* & Q_{22} \end{pmatrix}.$$

From (5) we obtain

$$B_1{}^*M_{11} + M_{11}B_1 = Q_{11}, \qquad B_2{}^*M_{22} + M_{22}B_2 = Q_{22}.$$

All eigenvalues of B_1 have positive real parts and $Q_{11} > 0$. From Ljapunov's Theorem we deduce $M_{11} > 0$ and similarly $M_{22} < 0$. Theorem 4 yields

$$\operatorname{In}(M) = \operatorname{In}\begin{pmatrix} M_{11} & 0 \\ 0 & M_{22} \end{pmatrix}, \qquad \operatorname{In}\begin{pmatrix} M_{11} & 0 \\ 0 & M_{22} \end{pmatrix} = \operatorname{In}(B) = \operatorname{In}(A)$$

and completes the first part of the proof.

(b) Theorem 3 $\Rightarrow$ Theorem 4. Let G be a matrix as in (3). We put

$$V := \begin{pmatrix} I_p & 0 \\ 0 & -I_n \end{pmatrix},$$

where I_p and I_n are $p \times p$ and $n \times n$ identity matrices.

$$G^*V + VG = 2\begin{pmatrix} G_{11} & 0 \\ 0 & -G_{22} \end{pmatrix} > 0.$$

Because of Theorem 3

$$\operatorname{In}(G) = \operatorname{In}(V) = (p, n, 0).$$

THEOREM 7. *If H is a Hermitian solution of*

$$H - A^*HA = C, \qquad C > 0,$$

then H is nonsingular and the number of positive (negative) eigenvalues of H is equal to the number of eigenvalues of A inside (outside) the unit circle.

Proof. The method and the notations are the same as in the proof of Theorem 6. Let λ be an eigenvalue of A and u an eigenvector corresponding to λ, $Au = \lambda u$. Then $u^*Hu - u^*A^*HAu = u^*Cu > 0$ and $u^*Hu(1 - \bar{\lambda}\lambda) > 0$. Thus $|\lambda| \neq 1$ for any eigenvalue λ of A and the Jordan form B of A can be partitioned into B_1 and $B_2 : B_1$ containing the blocks corresponding to eigenvalues with modulus less than 1 and B_2 with blocks corresponding to eigenvalues of A outside the unit circle. We get the equations

$$M_{11} - B_1{}^*M_{11}B_1 = Q_{11} \qquad \text{and} \qquad M_{22} - B_2{}^*M_{22}B_2 = Q_{22}\,.$$

Theorem 2 yields $M_{11} > 0$ and $M_{22} < 0$. By Theorem 4 we conclude that the number of positive eigenvalues of M is equal to the order of M_{11} which is equal to the number of eigenvalues of B with modulus less than 1. That there was an analogon to Theorem 3 is mentioned in [3].

Proof of Theorem 5. If (2) holds, then

$$(AP_+)^*\, P_+{}^*HP_+ + P_+{}^*HP_+(AP_+) = P_+{}^*CP_+\,.$$

We regard AP_+, $P_+{}^*HP_+$ and $P_+{}^*HP_+$ as operators on $P_+\mathbf{C}^n$. All eigenvalues of AP have positive real parts, $P_+{}^*CP_+$ is positive definite. From Theorem 1 follows that $P_+{}^*HP_+$ is positive definite on $P_+\mathbf{C}^n$.

THEOREM 8. *An $n \times n$ matrix A has the property*

$$R > 0 \Rightarrow A^*R + RA > 0, \tag{6}$$

if and only if

$$A = \alpha I, \qquad \operatorname{Re}\alpha > 0. \tag{7}$$

Proof. It is obvious that any matrix (7) has the property (6). To show the converse, assume (6) holds and define $C := A^*R + RA$. $C > 0$ implies that each principal minor of C is positive. If R is diagonal, $R = \operatorname{diag}(r_j)$, $r_j > 0$, then

$$C = (c_{jk}) = (\overline{a_{kj}}r_k + r_ja_{jk}).$$

Let H_{jk} be a 2×2 principal minor of $|C|$

$$H_{jk} := \begin{vmatrix} c_{jj} & c_{jk} \\ c_{kj} & c_{kk} \end{vmatrix} = (a_{jj} + \overline{a_{jj}})(a_{kk} + \overline{a_{kk}}) r_j r_k - |\overline{a_{kj}} r_k + r_j a_{jk}|^2.$$

If $a_{kj} \neq 0$ or $a_{jk} \neq 0$, r_j or r_k can be found, such that $H_{jk} < 0$. Then C would not be positive definite. Therefore A has to be diagonal, $A = \operatorname{diag}(\alpha_j)$, $\operatorname{Re} \alpha_j > 0$. Suppose there are two different diagonal elements in A, e.g., $\alpha_1 \neq \alpha_2$. We put

$$\hat{A} := \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}, \qquad \hat{P} := \begin{pmatrix} 1 & s \\ s & 1 \end{pmatrix}.$$

For s real and $s < 1$ $\hat{P}$ is positive definite. If $\alpha_1 \neq \alpha_2$, we can find $s < 1$, such that $\hat{C} := \hat{A}^* \hat{P} + \hat{P} \hat{A}$ is not positive definite.

$$|\hat{C}| = (\alpha_1 + \overline{\alpha_1})(\alpha_2 + \overline{\alpha_2}) - s^2(\overline{\alpha_1} + \alpha_2)(\alpha_1 + \overline{\alpha_2})$$

$\alpha_1 \neq \alpha_2$ is equivalent to

$$(\alpha_1 + \overline{\alpha_1})(\alpha_2 + \overline{\alpha_2}) - (\overline{\alpha_1} + \alpha_2)(\alpha_1 + \overline{\alpha_2}) < 0.$$

By an appropriate choice of $s < 1$ $|\hat{C}|$ is negative and $\hat{C}$ is not positive definite. Now put $P := \hat{P} \oplus I_{n-2}$ with I_{n-2} as identity matrix of order $n - 2$. Then $P > 0$, but $A^*P + PA$ is not positive definite, which contradicts to the assumption on A.

If A is a real matrix and $A^T P + PA < 0$ for each positive definite symmetric matrix P, then A is a multiple of the identity matrix, $A = aI$ with a real and $a < 0$. This has an application to the stability theory of linear systems of differential equations.

COROLLARY. *Each positive definite real quadratic form is a Ljapunov function for the real system $\dot{x} = Ax$, if and only if A has the form $A = aI$, $a < 0$.*

THEOREM 9. *Assume A has no purely imaginary eigenvalues. Let W be the set of all Hermitian matrices H which are positive definite on $P_+\mathbf{C}^n$ and negative definite on $P_-\mathbf{C}^n$. Then*

$$A^*H + HA > 0 \qquad \text{for all } H \in W, \tag{8}$$

if and only if A has the form

$$A = kP_+ - \bar{k}P_-, \qquad \operatorname{Re} k > 0. \tag{9}$$

Proof. For each $H \in W$

$$x^*P_+{}^*HP_+x \geqslant 0, \qquad x^*P_-{}^*HP_-x \leqslant 0.$$

The equality sign holds only if $P_+x = 0$, resp. $P_-x = 0$. $x \neq 0$ implies $P_+x \neq 0$ or $P_-x \neq 0$ and $x^*[P_+{}^*HP_+ - P_-{}^*HP_-]\,x > 0$. Let A have the form (9) and take $H \in W$. Then for any $x \neq 0$

$$x^*(A^*H + HA)\,x = (k + \bar{k})\,x^*[P_+{}^*HP_+ - P_-{}^*HP_-]\,x > 0,$$

which is equivalent to $A^*H + HA > 0$.

To prove the converse, we assume (8). For $u \in P_+\mathbf{C}^n$ and $H \in W$

$$u^*[(AP_+)^*\,P_+{}^*HP_+ + P_+{}^*HP_+(AP_+)]\,u > 0. \tag{10}$$

The set of all Hermitian matrices which are positive definite on $P_+\mathbf{C}^n$ is equal to the set $\mathscr{F} := \{F \mid F = P_+{}^*HP_+\,,\, H \in W\}$. AP_+ maps $P_+\mathbf{C}^n$ into itself. The set of all positive definite operators of the space $P_+\mathbf{C}^n$ can be identified with the set of matrices $\mathscr{F}$. Equation (10) and Theorem 8 imply $AP_+ = k_1P_+$ and $\operatorname{Re} k_1 > 0$. In a similar way we obtain $AP_- = k_2P_-$, $\operatorname{Re} k_2 < 0$. Thus $A = k_1P_+ + k_2P_-$ and

$$\begin{aligned} x^*(A^*H + HA)\,x = x^*[(k_1 + \overline{k_1})\,P_+{}^*HP_+ + (k_2 + \overline{k_2})\,P_-{}^*HP_-]\,x \\ + \operatorname{Re}[(k_1 + \overline{k_2})\,x^*P_-{}^*HP_+x]. \end{aligned} \tag{11}$$

Whether H is in W or not, does not depend on $P_-{}^*HP_+$. If $k_1 + \overline{k_2} \neq 0$, then for an $x \neq 0$ a Hermitian matrix H can be found, such that the sign in (11) is equal to the sign of $\operatorname{Re}[(k_1 + \overline{k_2})\,x^*P_-{}^*HP_+x]$. For a suitably chosen $x \neq 0$ and $H \in W$ (11) is negative and $A^*H + HA$ is not positive definite, which is a contradiction. Therefore $k_1 + \overline{k_2} = 0$.

Acknowledgment

I am grateful to Prof. A. D. Ziebur for valuable remarks.

References

1. A. Ostrowski and H. Schneider, Some theorems on the inertia of general matrices, *J. Math. Analysis Appl.* **4** (1962), 72–84.
2. O. Taussky, A generalisation of a theorem of Lyapunov, *J. Soc. Industr. Appl. Math.* **9** (1961), 640–643.
3. O. Taussky, Matrices C with $C^n \to 0$, *J. Algebra* **1** (1964), 5–10.
4. A. D. Ziebur, On determining the structure of A by analysing e^{At}, *SIAM Rev.* **12** (1970), 98–102.