CIDE

NÚMERO 28

David Mayer Foulkes

A CHARACTERISTIC POLYNOMIAL FACTORIZATION: EXTENDING SOME MATRIX EIGENVALUE ESTIMATES

Abstract. For matrices $E \in \mathbb{C}^{n\times n}$, $G, P \in \mathbb{C}^{n\times m}$, $J \in \mathbb{C}^{m\times m}$, let

$$A = E - PG^T \in \mathbb{C}^{n\times n}, \quad \mathscr{E} = \begin{pmatrix} J - G^T P & G^T \\ PJ - EP & E \end{pmatrix} \in \mathbb{C}^{(n+m)\times(n+m)}$$

Then the characteristic polynomials of J, A and $\mathscr{E}$ are related by $p_J(z)p_A(z) = p_{\mathscr{E}}(z)$. Eigenvalue estimates for several types of matrices A may be obtained by examining $\mathscr{E}$. We estimate *a)* the eigenvalues of Mondriga (and other) mixed-sign matrices, which arise in the geometry of immobilization as posed by Kuperberg and Papadimitriou, and which motivated this factorization; *b)* the eigenvalues of non-negative matrices different from the spectral radius; *c)* the eigenvalues of matrices with one sign on and another off the main diagonal. Also, we show a method for shifting the Geršgorin Disks.

Introduction

In this article we present several matrix theorems. The first of these, the theorem on Mondriga matrices, originated in the study of the immobilization of plane and solid figures in Euclidean space by points on their boundary, as posed by Kuperberg and Papadimitriou. In this context the theorem provides the tool to prove that if a tetrahedron satisfies a first-order immobilization condition, it also satisfies a second-order immobilization condition [2]. The first proofs obtained involved the study of many cases (given by the faces of n-dimensional polyhedra). Then we found a characteristic polynomial factorization which gave the result simply. This factorization in turn yielded several seemingly unrelated theorems.

1. The Characteristic Polynomial Factorization

Let $\mathbb{C}^{n\times m}$ represent the set of matrices with n rows and m columns over the field of complex numbers. Write $p_M(z) = \det(zI - M)$ for the characteristic polynomial of any square matrix M, σ_M for its spectrum.

1.1. Theorem. Let $G, P \in \mathbb{C}^{n\times m}$, $E \in \mathbb{C}^{n\times n}$, and suppose

$$A = E - PG^T \in \mathbb{C}^{n\times n}. \tag{1.1.1}$$

For any $J \in \mathbb{C}^{m\times m}$, let $K = J - G^T P$ and

$$\mathscr{E} = \begin{pmatrix} J - G^T P & G^T \\ PJ - EP & E \end{pmatrix} = \begin{pmatrix} K & G^T \\ PK - AP & A + PG^T \end{pmatrix} \in \mathbb{C}^{(n+m)\times(n+m)}. \tag{1.1.2}$$

Then

$$p_J(z)p_A(z) = p_{\mathscr{E}}(z), \tag{1.1.3}$$

$$\sigma_A \cup \sigma_J = \sigma_{\mathscr{E}}. \tag{1.1.4}$$

Thus the eigenvalues of A are those of $\mathscr{E}$ minus those of J.

Proof. The following identities may be directly verified:

$$\mathscr{E} = \begin{pmatrix} I & G^T \\ P & E \end{pmatrix}\begin{pmatrix} J & 0 \\ -P & I \end{pmatrix}; \qquad \begin{pmatrix} J & JG^T \\ 0 & A \end{pmatrix} = \begin{pmatrix} J & 0 \\ -P & I \end{pmatrix}\begin{pmatrix} I & G^T \\ P & E \end{pmatrix}.$$

Since $p_{MN}(z) = p_{NM}(z)$ for any matrices M, N, we obtain equation 1.1.3.■

Now suppose we are given a matrix A. By suitably choosing G, P, and J (or K), we may estimate the eigenvalues of A by estimating those of $\mathscr{E}$. The remaining sections of this note give several such applications to cases of interest.

2. Mondriga Matrices

The matrices naturally appearing in the second order condition for the immobilization of bodies in space, are the following (see [2]).

2.1. Definition. A Mondriga matrix is a matrix $A = (a_{il}) \in \mathbb{R}^{n \times n}$ whose diagonal entries are non-negative and the largest entry in each column, and the sum of whose rows in non-negative, i.e.,

$$\Sigma_l\, a_{il} \geq 0, \;\; a_{il} \leq a_{ll}, \;\; a_{ll} \geq 0,$$

where $1 \leq i, l \leq n$.■

Write $B_R(C) = \{z \in \mathbb{C} : | z - C | \leq R\}$ for the disk in the complex plane with radius R and center C.

2.2. Theorem. For a Mondriga matrix A, $\sigma_A \subset B_{\mathrm{trA}}(0)$.

Proof. Using the notation of theorem 1.1, set $m = 1$,

$$P = (1 \,, \ldots, 1)^T, \quad G = (a_{11} \,, \ldots, a_{nn})^T, \quad E = a_{ll} - a_{il} \geq 0, \quad J = (\mathrm{trA}) \,.$$

Then $- A = E - PG^T$ and

$$G^T P = J, \quad (PJ - EP)_i = trA - \Sigma_l (a_{ll} - a_{il}) = \Sigma_l a_{il} \geq 0 .$$

Hence by theorem 1.1 $\mathscr{E} = \begin{pmatrix} 0 & G^T \\ PJ - EP & E \end{pmatrix}$ is a non-negative matrix each of whose diagonal entries is zero, and each of whose rows has sum not exceeding trA. Thus applying either the Geršgorin Disk Theorem or the Frobenius Theorem we obtain $\sigma_{\mathscr{E}} \subset B_{trA}(0)$. Hence

$$\sigma_{-A} \subset \sigma_{\mathscr{E}} \subset B_{trA}(0) \;\text{ so }\; \sigma_A \subset B_{trA}(0) . \blacksquare$$

The extra eigenvalue introduced by $\mathscr{E}$ is trA, which coincides with the positive eigenvalue given by the Perron-Frobenius theorem, which therefore gives no additional information about A.

3. Applications to Mixed-Sign Real Matrices

The result on Mondriga Matrices given above may be refined as the corollary of a more general theorem.

First, we use a general vector of weights $\mathbf{p} > 0$. To apply the decomposition of theorem 1.1, given a matrix $A = (a_{il}) \in \mathbb{R}^{n \times n}$, we shall define $\mathbf{g} \geq 0$ so that

$$E = (e_{il}) = A + \boldsymbol{p}\boldsymbol{g}^T \geq 0 . \tag{3.1.1}$$

Thus

$$g_l = \max_{1 \leq i \leq n} \{ \max (-p_i^{-1} a_{il}), 0 \} \geq 0 . \tag{3.1.2}$$

Also define the quantities

$$\begin{gathered} \gamma = \mathbf{g}^T \mathbf{p}, \quad e = \min \{e_{11}, \ldots, e_{nn}\}, \\ j_0 = \max_{1 \leq i \leq n} p_i^{-1} (E\mathbf{p})_1 \geq e, \quad k_0 = \max_{1 \leq i \leq n} p_i^{-1} (A\mathbf{p})_i = j_0 - \gamma \end{gathered} \tag{3.1.3}$$

Observe that E, e, γ, j_0, k_0, are functions of the weights $\mathbf{p}$ which are independent of the norm $|\, \mathbf{p} \,|$.

3.1. Theorem. Let $A = (a_{il}) \in \mathbb{R}^{n \times n}$ be any matrix. For each $\mathbf{p} > 0$

$$\sigma_A \subset \begin{cases} B_{k_0+\gamma-e}(e) & e \le k_0 \\ B_\gamma(k_0) \cap B_\gamma(e) & e \ge k_0 \end{cases} \tag{3.1.4}$$

(where k_0, e, γ depend on $\mathbf{p}$).

Proof. Chose any $j \ge j_0$. By applying theorem 2.2 with $G = \mathbf{g}$, $P = \mathbf{p}$ we see that except for j, A has the eigenvalues of

$$\mathscr{C} = \begin{pmatrix} j-\gamma & \mathbf{g}^{\mathbf{T}} \\ j\mathbf{p} - \mathbf{Ep} & \mathrm{E} \end{pmatrix},$$

which is a matrix with non-negative entries except possibly for $j-\gamma$. Applying the Geršgorin Disk Theorem with weights $(1, \mathbf{p})$ (see [5]) we obtain that the eigenvalues of $\mathscr{C}$ lie on the disks

$$\mathrm{B}_{j-e_{ii}}(e_{ii}), \quad i = 1, \ldots, n, \quad \mathrm{B}_\gamma(j-\gamma)$$

since

$$\rho_i = \left| j - p_i^{-1} \sum_{l=1}^{n} p_l e_{il} \right| + p_i^{-1} \sum_{\substack{l=1 \\ l \ne i}}^{n} p_l e_{il} = j - e_{ii}.$$

Each of the first set of disks is contained in the largest, which has center at e and radius $j-e$. Hence

$$\sigma_A \subset \bigcap_{j \ge j_0} \left(\mathrm{B}_{j-e}(e) \cup \mathrm{B}_\gamma(j-\gamma) \right).$$

We have a pair of circunferences for each j, whose union is

$$\mathrm{B}_{j-e}(e) \cup \mathrm{B}_\gamma(j-\gamma) = \begin{cases} \mathrm{B}_{j-e}(e) & j \ge \gamma + e \\ \mathrm{B}_\gamma(j-\gamma) & j \le \gamma + e \end{cases}$$

Let $\mathrm{I}_1 = \{ j : j_0 \le j \le \gamma + e \}$ (empty if $\gamma + e < j_0$), and let $\mathrm{I}_2 = \left(\max \{ j_0, \gamma + e \}, \infty \right)$. Thus,

$$\sigma_A \subset \bigcap_{j \in \mathrm{I}_1} \mathrm{B}_\gamma(l-\gamma) \cap \bigcap_{j \in \mathrm{I}_2} \mathrm{B}_{j-e}(e)$$

Therefore if $\gamma + e \le j_0$

$$\sigma_A \subset \bigcap_{j \in I_2} B_{j-e}(e) = B_{j_0-e}(e)$$

while if $\gamma + e \geq j_0$

$$\sigma_A \subset B_\gamma(j_0-\gamma) \cap B_\gamma(e) \cap B_{j_0-e}(e) \subset B_\gamma(j_0-\gamma) \cap B_\gamma(e) . \blacksquare$$

The following corollary of the theorem 3.1 refines the theorem on Mondriga matrices presented above.

3.2. Corollary. Let $A = (a_{il}) \in \mathbb{R}^{n\times n}$ be a matrix satisfying, for some $\mathbf{p} > 0$, $A\mathbf{p} \geq 0$. If also

$$p_i^{-1} a_{il} \leq p_l^{-1} a_{ll} \quad \text{and} \quad p_l^{-1} a_{ll} \geq 0, \quad 1 \leq i\,, l \leq n, \tag{3.1.5}$$

$$\sigma_A \subset B_{trA}(\delta) \cap B_{trA}(0)\,,$$

where $\delta = \min\limits_{1\leq i\leq n} \; p_i^{-1} (A\mathbf{p})_i \geq 0$.

Proof. Apply theorem 3.1 to –A, using

$$k_0 = \max_{1\leq i\leq n} \; p_i^{-1} (-A\mathbf{p})_i \leq 0\,, \quad g_l = \max_{1\leq i\leq n} \; \max\left\{(p_i^{-1} a_{il}) \,, 0\right\} = p_l^{\,-1} a_{il}\,,$$

$$\gamma = trA\,, \quad e_{ii} = 0, \quad e = 0 . \blacksquare$$

4. Application to Non-negative Matrices

For a non-negative matrix A we may obtain from theorem 3.1 the Perron-Frobenius Theorem estimate. In this case, for any $\mathbf{p} > 0, \mathbf{g} = 0$ so $E = A, \gamma = 0$, $j_0 = k_0 = \max\left\{p_i^{-1} (A\mathbf{p})_i : 1 \leq i \leq n\right\}$, and $e = \min(a_{ii}) \leq k_0$. From theorem 3.1,

$$\sigma_A \subset \bigcap_{\mathbf{p}>0} B_{k_0-e}(e) = B_{r-e}(e), \quad \text{where } r = \inf_{\mathbf{p}>0} k_0\,,$$

which is the estimate obtained by applying the Perron-Frobenius theorem (see [5]) to $A - eI$.

Let us now obtain an estimate for the remaining eigenvalues.

4.1. Theorem. Suppose that $A = (a_{il}) \in \mathbb{R}^{n\times n}$ is a non-negative irreductible matrix with spectral radius r corresponding to a positive right eigenvector $\mathbf{p} > 0$. Define the maximum vector $\mathbf{g} \geq 0$ for which $E = (e_{il}) = A - \mathbf{p}\mathbf{g}^T \geq 0$, that is,

$$\mathbf{g} = (g_l) \in \mathbb{R}^n \quad \text{by} \quad g_l = \min_{1 \le i \le n} p_i^{-1} a_{il} \ge 0, \quad l = 1, \ldots, n.$$

Let $j = r - \mathbf{g}^{\mathrm{T}}\mathbf{p}$, $e = \min\{e_{11}, \ldots, e_{nn}\}$. Then the eigenvalues of A other than r lie on the disk

$$\sigma_{\mathrm{A}} \setminus \{r\} \subset \mathrm{B}_{j \quad e}(e).$$

Proof. We have $0 \le \mathrm{E}\mathbf{p} = \mathrm{A}\mathbf{p} - \mathbf{p}\mathbf{g}^{\mathrm{T}}\mathbf{p} = j\mathbf{p}$ implies $j > 0$. Now apply theorem 1.1 setting $m = 1$, $\mathrm{P} = -\mathbf{p} \in \mathbb{C}^{n\times 1}$, $\mathrm{G} = \mathbf{g} \in \mathbb{C}^{n \times 1}$. The theorem implies that except for j the eigenvalues of A are the eigenvalues of $\mathscr{E}$, where

$$\mathscr{E} = \begin{pmatrix} j - \mathrm{G}^{\mathrm{T}}\mathrm{P} & \mathrm{G}^{\mathrm{T}} \\ \mathrm{P}j - \mathrm{EP} & \mathrm{E} \end{pmatrix} = \begin{pmatrix} r & -\mathbf{g}^{\mathrm{T}} \\ 0 & \mathrm{E} \end{pmatrix}.$$

Thus the eigenvalues of A are r and those of E except for $a\, j$. Applying the Geršgorin Disk with weights $\mathbf{p}$ to the matrix E, we find that its eigenvalues lie on the disks

$$\{z \in \mathbb{C} \,:\, |z - e_{ii}| \le \rho_i\}, \quad i = 1, \ldots, n,$$

where $\rho_i = p_i^{-1} \Sigma_{l \ne i}\, p_l\, e_{il} = j - e_{il} \ge 0$. But each of these disks is contained in the largest, which has center e and radius $j - e$.■

5. Matrices with Non-negative Diagonal, Non-positive Elsewhere

For another application we consider matrices having one sign on the diagonal and the opposite elsewhere, which is a case in which γ can be calculated and which will yield a result on M-matrices.

5.1. Theorem. Suppose $\mathrm{A} = (a_{il}) \in \mathbb{R}^{n \times n}$ satisfies

$$a_{ii} \ge 0, \quad a_{il} \le 0 \ \text{ for } \ i \ne l.$$

Let r be the spectral radius of the non-negative matrix $\alpha\mathrm{I} - \mathrm{A}$, where $\alpha = \sup_{1 \le i \le n} a_{ii}$. Then A has eigenvalue $\alpha - r$ of multiplicity one with non-negative eigenvector and

$$\sigma_{\mathrm{A}} \setminus \{\alpha - r\} \subset \mathrm{B}_{j-e}(\alpha - e) \cap \mathrm{B}_{r + \mathrm{tr}\mathrm{A} - \alpha}(0) \subset \mathrm{B}_r(\alpha) \cap \mathrm{B}_{r+\mathrm{tr}\mathrm{A}-\alpha}(0)$$

where j and e are the (non-negative) quantities defined in theorem 4.1 for the matrix $\alpha\mathrm{I} - \mathrm{A}$.

Proof. Using the notation in theorem 3.1 as applied to –A, for each $\mathbf{p} > 0$, $\mathbf{g}$ is given by $g_i = p_i^{-1} a_{ii}$. Therefore $\gamma = \mathbf{g}^{\mathbf{T}}\mathbf{p} = \mathrm{trA}$. But $\mathrm{E}\mathbf{p} = (\mathrm{trAI} - \mathrm{A})\,\mathbf{p}$, so

$$s \equiv \inf_{\mathbf{p}>0} \max_{1 \le i \le n} p_i^{-1}(\mathrm{E}(\mathbf{p})\mathbf{p})_i = \mathrm{trA} - \alpha + \inf_{\mathbf{p}>0} \max_{1 \le i \le n} p_i^{-1}((\alpha\mathrm{I} - \mathrm{A})\mathbf{p})_i$$

$$= \mathrm{trA} - \alpha + r,$$

Fix $\mathbf{p}$ as the non-negative eigenvector of $\alpha\mathrm{I} = \mathrm{A}$. Then

$$\mathrm{E}\mathbf{p} = s\mathbf{p}, \quad \mathrm{A}\mathbf{p} = (\mathrm{trA} - s)\mathbf{p}.$$

$\mathrm{E} = \mathbf{p}\mathbf{g}^{\mathbf{T}} - \mathrm{A}$ is irreductible since A is, so s coincides with the spectral radius given by the Perron-Frobenius theorem (pre-multiply by the left-eigenvector). Applying theorem 1 with $j = s$ shows that except for an s, –A has the eigenvalues of

$$\mathscr{E} = \begin{pmatrix} s - \mathrm{trA} & \mathbf{g}^{\mathrm{T}} \\ 0 & \mathrm{E} \end{pmatrix}.$$

Therefore

$$\sigma_{-\mathrm{A}} \backslash \{s - \mathrm{trA}\} = \sigma_{\mathrm{E}} \backslash \{s\} \subset \mathrm{B}_s(0).$$

On the other hand, applying theorem 4 to $\alpha\mathrm{I} - \mathrm{A}$ we obtain

$$\sigma_{-\mathrm{A}} \backslash \{r - \alpha\} = \left(\sigma_{\alpha\mathrm{I} - \mathrm{A}} \backslash \{r - \alpha\}\right) - \mathrm{trA} \subset \mathrm{B}_{j-e}(e - \alpha). \blacksquare$$

Matrices with non-negative diagonal, non-positive elsewhere, are called M-matrices if their inverse is non-negative.

5.2. Corollary. In terms of the notation above, if

$$\{\alpha - r\} \cup \left(\mathrm{B}_{j-e}(\alpha - e) \cap \mathrm{B}_{r+\mathrm{trA}-\alpha}(0)\right) \subset \mathrm{Int}\,(\mathrm{B}_{\mathrm{R}}(0)),$$

where $\mathrm{R} > 0$, then $\mathrm{RI} - \mathrm{A}$ is an M-matrix.

Proof. This is a consequence of a well-know theorem (see Theorem 2, §15.2 of [7]).■

6. Shifting the Geršgorin Disks

Suppose that we are given a matrix A for which we know part of the Jordan form. Then the spectrum of A may be shown to be contained in shifted Geršgorin disks.

6.1. Theorem. Let $A \in \mathbb{C}^{n\times n}$, and suppose $A\mathcal{P} = \mathcal{P}\mathcal{J}$ where $\mathcal{P} \in \mathbb{C}^{n\times m}$ and $\mathcal{J} \in \mathbb{C}^{m\times m}$ is a Jordan block. For any $G \in \mathbb{C}^{n\times m}$, $\mathbf{q} > 0$,

$$\sigma_A \setminus \sigma_{\mathcal{J}} \subset \bigcup_{i=1}^{n} B_{\rho_i(G)}(e_{ii}(G))$$

where $\rho_i(G) = q_i^{-1} \sum_{\substack{l=1 \\ l\neq i}}^{n} q_l |e_{il}|$, $E(G) = (e_{il}(G)) = A + PG^T$.

Proof. In theorem 1.1 let $P = \mathcal{P}$, $J = \mathcal{J} + G^TP$, $E = A + PG^T$.

Then $EP = (A + PG^T)P = P(\mathcal{J} + G^TP) = PJ$. Hence $\mathcal{E} = \begin{pmatrix} J & G^T \\ 0 & E \end{pmatrix}$, so $\sigma_A \setminus \sigma_{\mathcal{J}} \subseteq \sigma_E$. Apply the Geršgorin Disk theorem with weights $\mathbf{q}$.■

References

1. J. Bracho, L. Montejano and J. Urrutia, *Immobilization of Smooth Convex Figures*, preprint, 1992.
2. J. Bracho, H. Fetter, D. Mayer, L. Montejano, *Immobilization of Solids And Mondriga Quadratic Forms*, México, CIDE (Working Paper 18, E), 1993. To appear in the *Journal of the London Mathematical Society*.
3. J. Czyzowicz, I. Stojmenovic and Jorge Urrutia, *Immobilizing a Shape*, Hull Department d'Informatique, Université du Québec à Hull, Report RR90/11-18, 1990.
4. D. Hilbert and S. Cohn-Vossen, *Geometry and Imagination*, New York, Chelsea Publishing Company, 1952.
5. R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge, Cambridge University Press, 1991.
6. W. Kuperberg, *DIMACS Workshop on Polytopes*, Rutgers University, January, 1990.
7. Peter Lancaster and Miron Tismenetsky, *The Theory of Matrices*, Second Edition, Orlando, Florida, Academic Press.
8. X. Markenscoff, L. Ni and Ch. H. Papadimitriou, *The Geometry of Grasping*, Int. J. Robotics Research, 9, 1, 1990, pp. 61-74.