Gower, J. C. 1980. A modified Leverrier-Faddeev algorithm for matrices with multiple-eigenvalues. *Linear Algebra and its Applications.* 31 (JUN), pp. 61-70.

A Modified Leverrier-Faddeev Algorithm for Matrices with Multiple Eigenvalues

J. C. Gower
Rothamsted Experimental Station
Harpenden, Hertfordshire, AL5 2JQ, United Kingdom

Submitted by Gene H. Golub

ABSTRACT

The Leverrier algorithm as modified by Faddeev gives the characteristic equation of a matrix **A**, its inverse, and the eigenvector corresponding to a simple eigenvalue λ of **A**. These results are extended (1) to give a generalized inverse when **A** is not of full rank and (2) to examine the modification required when λ is a multiple eigenvalue.

1. INTRODUCTION

Suppose **A** is a square matrix of order n, with characteristic polynomial

$$p(\lambda)=\lambda^{n}+p_1\lambda^{n-1}+p_2\lambda^{n-2}+\cdots+p_n, \tag{1}$$

where $p(\lambda)=0$ has roots λ_i $(i=1,2,\dots,n)$. Leverrier's method [1] for computing $p(\lambda)$ is to calculate $s_r=\sum_{i=1}^{n}\lambda_i^r=\mathrm{Tr}(\mathbf{A}^r)$ and use Newton's formulae:

$$\begin{aligned} s_1 &= -p_1,\\ s_2+p_1s_1 &= -2p_2,\\ s_3+p_1s_2+p_2s_1 &= -3p_3,\\ &\cdots\cdots\cdots\\ s_n+p_1s_{n-1}+p_2s_{n-2}+\cdots+p_{n-1}s_1 &= -np_n \end{aligned} \tag{2}$$

to derive the coefficients p_i. D. K. Faddeev [2, p. 260] modified Leverrier's

 0024-3795/80/030061+10$01.75

method by considering the sequence

$$\begin{aligned}
&\mathbf{Y}_0=\mathbf{A}, && \\
&\mathbf{Y}_1=\mathbf{A}\mathbf{Y}_0+p_1\mathbf{A}, &\text{where}\quad & p_1=-Tr(\mathbf{Y}_0),\\
&\mathbf{Y}_2=\mathbf{A}\mathbf{Y}_1+p_2\mathbf{A}, && p_2=-\tfrac{1}{2}\mathrm{Tr}(\mathbf{Y}_1),\\
&\mathbf{Y}_3=\mathbf{A}\mathbf{Y}_2+p_3\mathbf{A}, && p_3=-\tfrac{1}{3}\mathrm{Tr}(\mathbf{Y}_2),\\
&\cdots\cdots\cdots\cdots\cdots\cdots && \cdots\cdots\cdots\cdots\\
&\mathbf{Y}_n=\mathbf{A}\mathbf{Y}_{n-1}+p_n\mathbf{A}, && p_n=-\frac{1}{n}\mathrm{Tr}(\mathbf{Y}_{n-1}).
\end{aligned} \tag{3}$$

The coefficients p_i obtained in (3) satisfy Newton's formulae (2) and therefore are precisely those of $p(\lambda)$. Further $\mathbf{Y}_n=0$, so that the sequence (3) terminates naturally at the nth step, if not before, and provided $\mathbf{A}$ is nonsingular, $A^{-1}=-(1/p_n)(\mathbf{Y}_{n-2}+p_{n-1}\mathbf{I})$. Faddeev and Faddeeva [2] state (without proof) that when the roots λ_i are distinct, then for any root λ

$$Y=\sum_{i=0}^{n-1}\lambda^{n-i-1}\mathbf{Y}_i \tag{4}$$

is a nonnull matrix all of whose columns satisfy the eigenvector relationship $\mathbf{AY}=\lambda\mathbf{Y}$. It follows that since the eigenvalues of $\mathbf{A}$ are distinct, then $\mathrm{rank}(\mathbf{Y})=1$ and every nonnull column of $\mathbf{Y}$ is a multiple of the unique right eigenvector corresponding to λ. Because $\mathbf{Y}$ is a polynomial in $\mathbf{A}$ we have $\mathbf{AY}=\mathbf{YA}$, so that every row of $\mathbf{Y}$ is a multiple of the unique left eigenvector corresponding to λ. This duality relationship is not further discussed, but carries through in all the following, where the term eigenvector is to be understood as referring to a right eigenvector.

The algorithm (3) requires $O(n^4)$ multiplications to obtain an inverse and is subject to unacceptable rounding errors. There are many superior numerical eigenvector algorithms. *The Leverrier-Faddeev algorithm is inefficient and inaccurate and is clearly unsuitable for numerical work.* However, the sequence (3) does have algebraic interest, for (4) exhibits the eigenvectors of $\mathbf{A}$ in an explicit polynomial form. When $\mathbf{A}$ is patterned or has some regular structure that generates matrices Y_i that also exhibits structure, then (4) may have an especially simple form. For example, Gower [3] discusses the simplifications that occur when $\mathbf{A}$ is skew-symmetric and further simplifications when $\mathbf{A}$ is a special type of skew-symmetric matrix. The study of patterned matrices interests statisticians concerned with the possibility of approximating observed matrices with irregular structure by theoretically derived matrices with regular structure. Patterned matrices tend to have some, at least, of their eigenvalues occurring more than once. When $\mathbf{A}$ is skew-symmetric, $\mathbf{A}^2$ is symmetric with eigenvalues occurring in equal pairs. Thus before (4) can be used with patterned matrices it is essential to analyze what modifications are required when eigenvalues are not distinct. This paper shows that repeated eigenvalues can usually, but not always, be accommodated in a modification of the matrix $\mathbf{Y}$. Fortunately the exceptions occur only in well-defined pathological cases that are unlikely to be of practical importance. The modified algorithm is developed in Sec. 3, incidentally providing a proof of Faddeev's statement concerning distinct eigenvalues, but first an isolated result on generalized inverses of $\mathbf{A}$ is established.

2. GENERALIZED INVERSE

THEOREM 1. *If $\mathbf{A}$ is of rank r and $p_r \neq 0$, then*

$$\mathbf{A}^- = -\frac{1}{p_r}\left(\mathbf{Y}_{r-2} - \frac{p_{r-1}}{p_r}\mathbf{Y}_{r-1}\right)$$

is a reflexive generalized inverse of $\mathbf{A}$.

Proof. The characteristic equation becomes

$$p(\lambda) = \lambda^{n-r}(\lambda^r + p_1\lambda^{r-1} + \cdots + p_r),$$

where each p_i can be zero, and the special form taken by the Cayley-Hamilton theorem is

$$\mathbf{A}^{r+1} + p_1\mathbf{A}^r + p_2\mathbf{A}^{r-1} + \cdots + p_r\mathbf{A} = 0.$$

Thus in (3) $\mathbf{Y}_r = 0$, so that the sequence now stops at the rth step, with the final two relationships

$$\begin{aligned} \mathbf{Y}_{r-1} &= \mathbf{A}\mathbf{Y}_{r-2} + p_{r-1}\mathbf{A}, \\ \mathbf{Y}_r &= 0 = \mathbf{A}\mathbf{Y}_{r-1} + p_r\mathbf{A}. \end{aligned} \tag{5}$$

It follows that when $p_r \neq 0$, then $\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}$.

From (5)

$$\mathbf{A}^{-}\mathbf{A}\mathbf{A}^{-}=\frac{1}{p_r^2}\left(\mathbf{Y}_{r-2}-\frac{p_{r-1}}{p_r}\mathbf{Y}_{r-1}\right)\mathbf{A}\left(\mathbf{Y}_{r-2}-\frac{p_{r-1}}{p_r}\mathbf{Y}_{r-1}\right)$$

$$=\frac{1}{p_r^2}\left(\mathbf{Y}_{r-2}-\frac{p_{r-1}}{p_r}\mathbf{Y}_{r-1}\right)\mathbf{Y}_{r-1}. \tag{6}$$

Now

$$\mathbf{Y}_{r-2}=\mathbf{A}^{r-1}+p_1\mathbf{A}^{r-2}+p_2\mathbf{A}^{r-3}+\cdots+p_{r-2}\mathbf{A},$$

$$\mathbf{Y}_{r-1}=\mathbf{A}^{r}+p_1\mathbf{A}^{r-1}+p_2\mathbf{A}^{r-2}+\cdots+p_{r-1}\mathbf{A},$$

$$\mathbf{Y}_{r}=\mathbf{A}^{r+1}+p_1\mathbf{A}^{r}+p_2\mathbf{A}^{r-1}+\cdots+p_r\mathbf{A}=\mathbf{O},$$

so that $\mathbf{Y}_{r-2}\mathbf{Y}_{r-1}=-p_r\mathbf{Y}_{r-2}$ and $\mathbf{Y}_{r-1}^2=-p_r\mathbf{Y}_{r-1}$ and (6) becomes

$$\mathbf{A}^{-}\mathbf{A}\mathbf{A}^{-}=\mathbf{A}^{-}. \qquad \blacksquare$$

Decell [4] expresses $\mathbf{A}^{+}$, the Moore-Penrose inverse of rectangular $\mathbf{A}$, in terms of the characteristic polynomial of $\mathbf{A}\mathbf{A}^*$, where $\mathbf{A}^*$ is the conjugate transpose of $\mathbf{A}$. The above gives a reflexive inverse of square $\mathbf{A}$ in terms of the characteristic polynomial of $\mathbf{A}$. Decell [4] does not require $p_r\neq 0$, but this restriction seems essential here for any generalized inverse that is a linear combination of the powers of $\mathbf{A}$, as is indicated by the following example:

Let

$$A=\begin{bmatrix}1&1&0\\0&0&1\\0&0&0\end{bmatrix},$$

which has characteristic polynomial $\lambda^2(\lambda-1)$. Thus $\mathbf{A}$ has rank 2, and $p_2=0$. We have

$$\mathbf{Y}_1=\mathbf{A}^2-\mathbf{A}=\begin{bmatrix}0&0&1\\0&0&-1\\0&0&0\end{bmatrix}\quad\text{and}\quad \mathbf{Y}_2=\mathbf{O}.$$

Now the middle element of the last row of every generalized inverse of $\mathbf{A}$ has $\mathbf{A}_{3,2}^{-}=1$, so $\mathbf{A}^{-}$ cannot be expressed as any linear combination of $\mathbf{Y}_0$, $\mathbf{Y}_1$, and $\mathbf{Y}_2$.

3. REPEATED EIGENVALUES

For $\mathbf{Y}$ given by (4) to give nonnull eigenvectors, Faddeev and Faddeeva [2, p. 264] require all eigenvalues to be distinct. Certainly without this condition $\mathbf{Y}$ may be null, as happens for example with $\mathbf{A}=\mathbf{I}$, but it is shown below that the condition is not a necessary one. In this section a modified procedure is described giving a matrix $\mathbf{X}$ of eigenvectors which is never null and whose rank, under suitable conditions, is equal to the multiplicity of the corresponding eigenvalue. The minimal rather than characteristic polynomial is the basis of the modifications. Let

$$q(\lambda)=\lambda^m+q_1\lambda^{m-1}+\cdots+q_m \tag{7}$$

be the minimal polynomial of $\mathbf{A}$, and consider the sequence

$$\begin{aligned}
&\mathbf{X}_0=\mathbf{A},\\
&\mathbf{X}_1=\mathbf{A}\mathbf{X}_0+q_1\mathbf{A},\\
&\mathbf{X}_2=\mathbf{A}\mathbf{X}_1+q_2\mathbf{A},\\
&\cdots\cdots\cdots\cdots\cdots\\
&\mathbf{X}_m=\mathbf{A}\mathbf{X}_{m-1}+q_m\mathbf{A}.
\end{aligned} \tag{8}$$

Let

$$\begin{aligned}
&\mathbf{X}\sum_{i=0}^{m-1}\lambda^{m-i-1}\mathbf{X}_i &&\text{when}\quad \lambda\neq 0\\
&\mathbf{X}=\mathbf{X}_{m-2}+q_{m-1}\mathbf{I} &&\text{when}\quad \lambda=0.
\end{aligned}$$

In this sequence the coefficients q_i are assumed known. I know of no modification, similar to that of Faddeev for the characteristic polynomial, that allows these coefficients to be found sequentially. Let k be the number of separate Jordan blocks corresponding to the root λ, and let l_λ be the number of these blocks that have maximum size.

THEOREM 2.

(i) $\mathbf{X}_m=\mathbf{O}$,
(ii) $\mathbf{A}\mathbf{X}=\lambda\mathbf{X}$, *where* $\mathbf{X}$ *is not null*,
(iii) $\operatorname{rank}(\mathbf{X})=l_\lambda$.

Proof. The first part follows immediately from $\mathbf{X}_m = \mathbf{A}_q(\mathbf{A}) = 0$, and, leaving aside for a while the case $\lambda = 0$, (ii) is a simple consequence of (7) and the definition of **X**. This shows that the columns of **X**, provided they are nonnull, are all eigenvectors of **A** corresponding to the root λ. Next we show that **X** *is nonnull.* Expanding **X** *in terms of* **A** gives

$$\mathbf{X} = \lambda^{m-1}\mathbf{A} + \lambda^{m-2}(\mathbf{A}^2 + q_1\mathbf{A}) + \lambda^{m-3}(\mathbf{A}^3 + q_1\mathbf{A}^2 + q_2\mathbf{A}) + \cdots$$
$$+ (\mathbf{A}^m + q_1\mathbf{A}^{m-1} + q_2\mathbf{A}^{m-2} + \cdots + q_{m-1}\mathbf{A}),$$

which is a polynomial in **A** of degree m, and so can only vanish if it coincides with the (unique) minimal polynomial. This requires

$$1 = \frac{q_1 + \lambda}{q_1} = \frac{q_2 + \lambda q_1 + \lambda^2}{q_2} = \cdots = \frac{q_{m-1} + \lambda q_{m-2} + \cdots + \lambda^{m-1}}{q_{m-1}}, \tag{9}$$

$$q_m = 0.$$

When $\lambda \neq 0$ (9) is impossible, so **X** is not null. When $\lambda = 0$ is a root, $q_m = 0$ [from (7)] and hence (9) is valid and **X**, defined as for $\lambda \neq 0$, is null. In this case we have

$$\mathbf{X}_{m-1} = \mathbf{A}^m + q_1\mathbf{A}^{m-1} + \cdots + q_{m-1}\mathbf{A} = -q_m\mathbf{I} = \mathbf{0},$$

so that the sequence terminates one step earlier than usual to give

$$\mathbf{A}(\mathbf{X}_{m-2} + q_{m-1}\mathbf{I}) = \mathbf{0},$$

showing that $\mathbf{X}_{m-2} + q_{m-1}\mathbf{I}$ has columns corresponding to the zero root. But $\mathbf{X}_{m-2} + q_{m-1}\mathbf{I}$ is a polynomial in **A** of degree $m-1$, and therefore cannot vanish. Thus whether or not λ is zero, **X** is not null and satisfies $\mathbf{AX} = \lambda\mathbf{X}$, completing the proof of part (ii) of the theorem.

To establish the rank of **X**, let $\mathbf{TAT}^{-1} = \mathbf{J}$ be the Jordan form of **A**; then

$$\mathbf{JZ} = \lambda\mathbf{Z},$$

where $\mathbf{Z} = \mathbf{TX}$, and because **T** is nonsingular, rank $(\mathbf{Z}) = \text{rank}(\mathbf{X})$. Consider a Jordan block, say $\mathbf{J}_0$, corresponding to a root μ. Suppose **J** is of order t, and let $\mathbf{Z}^*$ be the t rows of $\mathbf{Z}$ that are multiplied by $\mathbf{J}_0$. We have

$$\begin{bmatrix} \mu & 1 & & \\ & \mu & \ddots & \\ & & \ddots & 1 \\ & & & \mu \end{bmatrix} \mathbf{Z}^* = \lambda \mathbf{Z}^*,$$

so that

$$\begin{aligned} \mu \mathbf{z}_1 + \mathbf{z}_2 &= \lambda \mathbf{z}_1, \\ \mu \mathbf{z}_2 + \mathbf{z}_3 &= \lambda \mathbf{z}_2, \\ &\cdots\cdots\cdots \\ \mu \mathbf{z}_{t-1} + \mathbf{z}_t &= \lambda \mathbf{z}_{t-1}, \\ \mu \mathbf{z}_t &= \lambda \mathbf{z}_t, \end{aligned} \tag{10}$$

where $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_t$ are the rows of $\mathbf{Z}^*$.

When $\lambda \neq \mu$, (10) shows that $\mathbf{z}_1 = \mathbf{z}_2 = \cdots = \mathbf{z}_t = 0$. When $\lambda = \mu$, $\mathbf{z}_2 = \mathbf{z}_3 = \cdots = \mathbf{z}_t = 0$. Hence $\mathbf{Z}$ can have nonzero rows only for the first row of each Jordan block in $\mathbf{J}$ with the root λ.

Now $\mathbf{X}$ is a polynomial in $\mathbf{A}$ of degree $d = m$, unless $\lambda = 0$, in which case the polynomial has degree $d = m - 1$. Therefore

$$\mathbf{X} = \sum_{i=0}^{d} \gamma_i \mathbf{A}^i,$$

where

$$\begin{aligned} \gamma_0 &= 0, & \lambda &\neq 0 \\ &= q_{m-1} & \lambda &= 0. \end{aligned}$$

We have

$$\mathbf{Z} = \mathbf{T}\mathbf{X} = \sum_{i=0}^{d} \gamma_i (\mathbf{T}\mathbf{A}^i\mathbf{T}^{-1})\mathbf{T} = \left(\sum_{i=0}^{d} \gamma_i \mathbf{J}^i \right) \mathbf{T}. \tag{11}$$

It was shown above that the only possible nonzero rows of $\mathbf{Z}$ are those that correspond to the first row of each of the k_λ Jordan blocks with eigenvalue λ. Assume these blocks $\mathbf{J}_1, \mathbf{J}_2, \ldots, \mathbf{J}_{k_\lambda}$ are labeled in order of decreasing size $k_1 \geqslant k_2 \geqslant \cdots \geqslant k_{k_\lambda}$; then it is easily seen that $\mathbf{K} = \Sigma_{i=0}^{d} \gamma_i \mathbf{J}^i$ is made up of k_λ blocks with form

$$\mathbf{K}_r = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_{k_r} \\ & a_1 & a_2 & \cdots & a_{k_r} - 1 \\ & & a_1 & \cdots & a_{k_r} - 2 \\ & & & \ddots & \vdots \\ & & & & a_1 \end{bmatrix}$$

where

$$a_{j+1} = \sum_{i=j}^{d} \gamma_i \binom{i}{j} \lambda^{i-j}.$$

The largest such matrix is $\mathbf{K}_1$; subsequent smaller matrices in the series will omit the final rows and columns of $\mathbf{K}_1$. Only the first row of $\mathbf{Z}^*$ corresponding to $\mathbf{K}_1$ can be nonzero. Hence from (11)

$$\mathbf{K}_1 \begin{bmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \\ \vdots \\ \mathbf{t}_{k_1} \end{bmatrix} = \begin{bmatrix} \mathbf{z}_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

where $\mathbf{t}_i$ is the ith row of $\mathbf{T}$ that is multiplied by $\mathbf{K}_1$, and so

$$\begin{aligned} a_1\mathbf{t}_1 + a_2\mathbf{t}_2 + \cdots + a_{k_1}\mathbf{t}_{k_1} &= \mathbf{z}_1, \\ a_1\mathbf{t}_2 + \cdots + a_{k_1-1}\mathbf{t}_{k_1} &= 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ a_1\mathbf{t}_{k_1} &= 0. \end{aligned}$$

Since $\mathbf{T}$ is nonsingular, no row of $\mathbf{T}$ is null. In particular $\mathbf{t}_{k_1} \neq 0$ and hence $a_1 = a_2 = \cdots = a_{k_1-1} = 0$. If in addition $a_{k_1} = 0$, then $\mathbf{z}_1 = 0$, as are the first rows of each $\mathbf{Z}^*$ corresponding to the other Jordan blocks. This implies $\mathbf{Z} = 0$, which contradicts the previous result; hence $a_{k_1} \neq 0$. It follows that $\mathbf{K}_i = 0$ except for $\mathbf{K}_1$ and any blocks of the same size, which each have one nonzero element equal to a_{k_1}. By definition there are l_λ blocks of the same size as $\mathbf{K}_1$, so

$$\operatorname{rank}(\mathbf{K}) = \operatorname{rank}(\mathbf{X}) = l_\lambda. \qquad \blacksquare$$

It is instructive to examine the circumstances under which the proof of Theorem 2 breaks down when we work with $\mathbf{Y}$ derived from the characteristic polynomial rather than with $\mathbf{X}$ derived from the minimal polynomial. Firstly, when the characteristic polynomial differs from the minimal polynomial, (9) is invalid, so it does not follow that $\mathbf{Y}$ given by (4) is necessarily nonnull. This in turn admits the possibility that a_{k_1} may be zero. In these circumstances a_{k_1} must be examined in more detail. Writing s for $k_1 - 1$, we have

$$a_{k_1} = \sum_{i=s}^{n} \gamma_i \binom{i}{s} \gamma^{i-s},$$

where now γ_i is the coefficient of $\mathbf{A}^i$ in the expression for $\mathbf{Y}$. Thus

$$\gamma_i = \sum_{j=0}^{n-i} p_{n-i-j} \lambda^j$$

and

$$\begin{aligned} a_{k_1} &= \sum_{i=s}^{n} \left(\sum_{j=0}^{n-i} p_{n-i-j} \lambda^j \right) \binom{i}{s} \lambda^{i-s} \\ &= \sum_{i=0}^{n-s} p_i \lambda^{n-s-i} \left[\binom{s}{s} + \binom{s+1}{s} + \cdots + \binom{n-i}{s} \right] \\ &= \sum_{i=0}^{n-s} p_i \lambda^{n-s-i} \binom{n-i+1}{s+1}. \end{aligned} \tag{12}$$

When λ is an eigenvalue of greater multiplicity than $k_1 = s+1$, we can differentiate $\lambda p(\lambda)$ $s+1$ times, showing that (12) is zero and $\mathbf{Y}$ is null. However when the multiplicity of λ is k_1, differentiating $s+1$ times does not yield a zero polynomial in λ and (12) is not zero. Thus if an eigenvalue λ occurs in more than one Jordan block, $\mathbf{Y}$ is null. If λ occurs in a single Jordan block, $\mathbf{Y}$ has rank 1 and its columns are representations of the single eigenvector associated with λ.

Theorem 2 has shown that $\mathbf{X}$ derived via the minimal polynomial, unlike $\mathbf{Y}$ derived from the characteristic polynomial, is never null. However even $\mathbf{X}$ may not span the space of all independent eigenvectors corresponding to an eigenvalue λ; eigenvectors arising from submaximal Jordan blocks associated with λ will not be generated. When the eigenvalues are distinct, minimal and characteristic polynomials coincide and $l_\lambda = 1$; therefore $\mathbf{X} = \mathbf{Y}$ with rank 1, which is Faddeev's result. When $\mathbf{J}$ is diagonal, $\mathbf{X}$ has rank equal to the multiplicity of the root λ, and so gives all the vectors; in particular this includes the cases where $\mathbf{A}$ is symmetric, Hermitian, or skew-symmetric. However, when $\mathbf{J}$ is diagonal, $\mathbf{Y}$ is null for any multiple eigenvalue. Gower [3] applies these results to obtain explicit singular value decompositions of certain skew-symmetric matrices.

REFERENCES

1 U. J. J. Leverrier, Sur les variations séculaires des élements des orbites, *J. Math.*, 1840.
2 D. K. Faddeev and V. N. Faddeeva, *Computational Methods of Linear Algebra* (in Russian), Fizmatgiz, Moscow, 1960; English translation by R. C. Williams, W. H. Freeman, San Francisco, 1963.
3 J. C. Gower, An application of the Leverrier-Faddeev algorithm to skew-symmetric matrix decompositions, submitted for publication.
4 H. P. Decell, An application of the Cayley-Hamilton theorem to generalized matrix inversion, *SIAM Rev.*, 7:526–528 (1965).

Received 16 March 1979; revised 13 June 1979