WIKIPEDIA
The Free Encyclopedia

Eigenvalue algorithm

In numerical analysis, one of the most important problems is designing efficient and stable algorithms for finding the eigenvalues of a matrix. These **eigenvalue algorithms** may also find eigenvectors.

Eigenvalues and eigenvectors

Given an $n \times n$ square matrix A of real or complex numbers, an *eigenvalue* λ and its associated *generalized eigenvector* $\mathbf{v}$ are a pair obeying the relation[1]

$$(A - \lambda I)^k \mathbf{v} = 0,$$

where $\mathbf{v}$ is a nonzero $n \times 1$ column vector, I is the $n \times n$ identity matrix, k is a positive integer, and both λ and $\mathbf{v}$ are allowed to be complex even when A is real. When $k = 1$, the vector is called simply an *eigenvector*, and the pair is called an *eigenpair*. In this case, $A\mathbf{v} = \lambda\mathbf{v}$. Any eigenvalue λ of A has ordinary[note 1] eigenvectors associated to it, for if k is the smallest integer such that $(A - \lambda I)^k \mathbf{v} = 0$ for a generalized eigenvector $\mathbf{v}$, then $(A - \lambda I)^{k-1} \mathbf{v}$ is an ordinary eigenvector. The value k can always be taken as less than or equal to n. In particular, $(A - \lambda I)^n \mathbf{v} = 0$ for all generalized eigenvectors $\mathbf{v}$ associated with λ.

For each eigenvalue λ of A, the kernel $\ker(A - \lambda I)$ consists of all eigenvectors associated with λ (along with 0), called the *eigenspace* of λ, while the vector space $\ker((A - \lambda I)^n)$ consists of all generalized eigenvectors, and is called the *generalized eigenspace*. The *geometric multiplicity* of λ is the dimension of its eigenspace. The *algebraic multiplicity* of λ is the dimension of its generalized eigenspace. The latter terminology is justified by the equation

$$p_A(z) = \det(zI - A) = \prod_{i=1}^{k} (z - \lambda_i)^{\alpha_i},$$

where $\det$ is the determinant function, the λ_i are all the distinct eigenvalues of A and the α_i are the corresponding algebraic multiplicities. The function $p_A(z)$ is the *characteristic polynomial* of A. So the algebraic multiplicity is the multiplicity of the eigenvalue as a zero of the characteristic polynomial. Since any eigenvector is also a generalized eigenvector, the geometric multiplicity is less than or equal to the algebraic multiplicity. The algebraic multiplicities sum up to n, the degree of the characteristic polynomial. The equation $p_A(z) = 0$ is called the *characteristic equation*, as its roots are exactly the eigenvalues of A. By the Cayley–Hamilton theorem, A itself obeys the same equation: $p_A(A) = 0$.[note 2] As a consequence, the columns of the matrix $\prod_{i \neq j} (A - \lambda_i I)^{\alpha_i}$ must be either 0 or generalized eigenvectors of the eigenvalue λ_j, since they are annihilated by $(A - \lambda_j I)^{\alpha_j}$. In fact, the column space is the generalized eigenspace of λ_j.

Any collection of generalized eigenvectors of distinct eigenvalues is linearly independent, so a basis for all of $\mathbf{C}^n$ can be chosen consisting of generalized eigenvectors. More particularly, this basis $\{\mathbf{v}_i\}_{i=1}^{n}$ can be chosen and organized so that

- if $\mathbf{v}_i$ and $\mathbf{v}_j$ have the same eigenvalue, then so does $\mathbf{v}_k$ for each k between i and j, and
- if $\mathbf{v}_i$ is not an ordinary eigenvector, and if λ_i is its eigenvalue, then $(A - \lambda_i I)\mathbf{v}_i = \mathbf{v}_{i-1}$ (in particular, $\mathbf{v}_1$ must be an ordinary eigenvector).

If these basis vectors are placed as the column vectors of a matrix $V = [\mathbf{v}_1\ \mathbf{v}_2\ \cdots\ \mathbf{v}_n]$, then V can be used to convert A to its Jordan normal form:

$$V^{-1}AV = \begin{bmatrix} \lambda_1 & \beta_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \beta_2 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix},$$

where the λ_i are the eigenvalues, $\beta_i = 1$ if $(A - \lambda_{i+1})\mathbf{v}_{i+1} = \mathbf{v}_i$ and $\beta_i = 0$ otherwise.

More generally, if W is any invertible matrix, and λ is an eigenvalue of A with generalized eigenvector $\mathbf{v}$, then $(W^{-1}AW - \lambda I)^k\ W^{-k}\mathbf{v} = 0$. Thus λ is an eigenvalue of $W^{-1}AW$ with generalized eigenvector $W^{-k}\mathbf{v}$. That is, similar matrices have the same eigenvalues.

Normal, Hermitian, and real-symmetric matrices

The adjoint M^* of a complex matrix M is the transpose of the conjugate of M: $M^* = \overline{M}^{\mathrm{T}}$. A square matrix A is called *normal* if it commutes with its adjoint: $A^*A = AA^*$. It is called *Hermitian* if it is equal to its adjoint: $A^* = A$. All Hermitian matrices are normal. If A has only real elements, then the adjoint is just the transpose, and A is Hermitian if and only if it is symmetric. When applied to column vectors, the adjoint can be used to define the canonical inner product on $\mathbf{C}^n$: $\mathbf{w} \cdot \mathbf{v} = \mathbf{w}^*\ \mathbf{v}$.[note 3] Normal, Hermitian, and real-symmetric matrices have several useful properties:

- Every generalized eigenvector of a normal matrix is an ordinary eigenvector.
- Any normal matrix is similar to a diagonal matrix, since its Jordan normal form is diagonal.
- Eigenvectors of distinct eigenvalues of a normal matrix are orthogonal.
- The null space and the image (or column space) of a normal matrix are orthogonal to each other.
- For any normal matrix A, $\mathbf{C}^n$ has an orthonormal basis consisting of eigenvectors of A. The corresponding matrix of eigenvectors is unitary.
- The eigenvalues of a Hermitian matrix are real, since $(\overline{\lambda} - \lambda)\mathbf{v} = (A^* - A)\mathbf{v} = (A - A)\mathbf{v} = 0$ for a non-zero eigenvector $\mathbf{v}$.
- If A is real, there is an orthonormal basis for $\mathbf{R}^n$ consisting of eigenvectors of A if and only if A is symmetric.

It is possible for a real or complex matrix to have all real eigenvalues without being Hermitian. For example, a real triangular matrix has its eigenvalues along its diagonal, but in general is not symmetric.

Condition number

Any problem of numeric calculation can be viewed as the evaluation of some function f for some input x. The condition number $\kappa(f, x)$ of the problem is the ratio of the relative error in the function's output to the relative error in the input, and varies with both the function and the input. The condition number describes how error grows during the calculation. Its base-10 logarithm tells how many fewer digits of accuracy exist in the result than existed in the input. The condition number is a best-case scenario. It reflects the instability built into the problem, regardless of how it is solved. No algorithm can ever produce more accurate results than indicated by the condition number, except by chance. However, a poorly designed algorithm may produce significantly worse results. For example, as mentioned below, the problem of finding eigenvalues for normal matrices is

always well-conditioned. However, the problem of finding the roots of a polynomial can be very ill-conditioned. Thus eigenvalue algorithms that work by finding the roots of the characteristic polynomial can be ill-conditioned even when the problem is not.

For the problem of solving the linear equation $A\mathbf{v} = \mathbf{b}$ where A is invertible, the matrix condition number $\kappa(A^{-1}, \mathbf{b})$ is given by $||A||_{\text{op}}||A^{-1}||_{\text{op}}$, where $\|\ \|_{\text{op}}$ is the operator norm subordinate to the normal Euclidean norm on $\mathbf{C}^n$. Since this number is independent of $\mathbf{b}$ and is the same for A and A^{-1}, it is usually just called the condition number $\kappa(A)$ of the matrix A. This value $\kappa(A)$ is also the absolute value of the ratio of the largest eigenvalue of A to its smallest. If A is unitary, then $||A||_{\text{op}} = ||A^{-1}||_{\text{op}} = 1$, so $\kappa(A) = 1$. For general matrices, the operator norm is often difficult to calculate. For this reason, other matrix norms are commonly used to estimate the condition number.

For the eigenvalue problem, Bauer and Fike proved that if λ is an eigenvalue for a diagonalizable $n \times n$ matrix A with eigenvector matrix V, then the absolute error in calculating λ is bounded by the product of $\kappa(V)$ and the absolute error in A.[2] As a result, the condition number for finding λ is $\kappa(\lambda, A) = \kappa(V) = ||V||_{\text{op}}\ ||V^{-1}||_{\text{op}}$. If A is normal, then V is unitary, and $\kappa(\lambda, A) = 1$. Thus the eigenvalue problem for all normal matrices is well-conditioned.

The condition number for the problem of finding the eigenspace of a normal matrix A corresponding to an eigenvalue λ has been shown to be inversely proportional to the minimum distance between λ and the other distinct eigenvalues of A.[3] In particular, the eigenspace problem for normal matrices is well-conditioned for isolated eigenvalues. When eigenvalues are not isolated, the best that can be hoped for is to identify the span of all eigenvectors of nearby eigenvalues.

Algorithms

The most reliable and most widely used algorithm for computing eigenvalues is John G. F. Francis' QR algorithm, considered one of the top ten algorithms of 20th century.[4]

Any monic polynomial is the characteristic polynomial of its companion matrix. Therefore, a general algorithm for finding eigenvalues could also be used to find the roots of polynomials. The Abel–Ruffini theorem shows that any such algorithm for dimensions greater than 4 must either be infinite, or involve functions of greater complexity than elementary arithmetic operations and fractional powers. For this reason algorithms that exactly calculate eigenvalues in a finite number of steps only exist for a few special classes of matrices. For general matrices, algorithms are iterative, producing better approximate solutions with each iteration.

Some algorithms produce every eigenvalue, others will produce a few, or only one. However, even the latter algorithms can be used to find all eigenvalues. Once an eigenvalue λ of a matrix A has been identified, it can be used to either direct the algorithm towards a different solution next time, or to reduce the problem to one that no longer has λ as a solution.

Redirection is usually accomplished by shifting: replacing A with $A - \mu I$ for some constant μ. The eigenvalue found for $A - \mu I$ must have μ added back in to get an eigenvalue for A. For example, for power iteration, $\mu = \lambda$. Power iteration finds the largest eigenvalue in absolute value, so even when λ is only an approximate eigenvalue, power iteration is unlikely to find it a second time. Conversely, inverse iteration based methods find the lowest eigenvalue, so μ is chosen well away from λ and hopefully closer to some other eigenvalue.

Reduction can be accomplished by restricting A to the column space of the matrix $A - \lambda I$, which A carries to itself. Since $A - \lambda I$ is singular, the column space is of lesser dimension. The eigenvalue algorithm can then be applied to the restricted matrix. This process can be repeated until all eigenvalues are found.

If an eigenvalue algorithm does not produce eigenvectors, a common practice is to use an inverse iteration based algorithm with μ set to a close approximation to the eigenvalue. This will quickly converge to the eigenvector of the closest eigenvalue to μ. For small matrices, an alternative is to look at the column space of the product of $A - \lambda' I$ for each of the other eigenvalues λ'.

A formula for the norm of unit eigenvector components of normal matrices was discovered by Robert Thompson in 1966 and rediscovered independently by several others.[5][6][7][8][9] If A is an $n \times n$ normal matrix with eigenvalues $\lambda_i(A)$ and corresponding unit eigenvectors $\mathbf{v}_i$ whose component entries are $v_{i,j}$, let A_j be the $n-1 \times n-1$ matrix obtained by removing the i-th row and column from A, and let $\lambda_k(A_j)$ be its k-th eigenvalue. Then

$$|v_{i,j}|^2 \prod_{k=1,k\neq i}^{n} (\lambda_i(A) - \lambda_k(A)) = \prod_{k=1}^{n-1} (\lambda_i(A) - \lambda_k(A_j))$$

If p, p_j are the characteristic polynomials of A and A_j, the formula can be re-written as

$$|v_{i,j}|^2 = \frac{p_j(\lambda_i(A))}{p'(\lambda_i(A))}$$

assuming the derivative p' is not zero at $\lambda_i(A)$.

Hessenberg and tridiagonal matrices

Because the eigenvalues of a triangular matrix are its diagonal elements, for general matrices there is no finite method like gaussian elimination to convert a matrix to triangular form while preserving eigenvalues. But it is possible to reach something close to triangular. An upper Hessenberg matrix is a square matrix for which all entries below the subdiagonal are zero. A lower Hessenberg matrix is one for which all entries above the superdiagonal are zero. Matrices that are both upper and lower Hessenberg are tridiagonal. Hessenberg and tridiagonal matrices are the starting points for many eigenvalue algorithms because the zero entries reduce the complexity of the problem. Several methods are commonly used to convert a general matrix into a Hessenberg matrix with the same eigenvalues. If the original matrix was symmetric or Hermitian, then the resulting matrix will be tridiagonal.

When only eigenvalues are needed, there is no need to calculate the similarity matrix, as the transformed matrix has the same eigenvalues. If eigenvectors are needed as well, the similarity matrix may be needed to transform the eigenvectors of the Hessenberg matrix back into eigenvectors of the original matrix.

Method	Applies to	Produces	Cost without similarity matrix	Cost with similarity matrix	Description
Householder transformations	General	Hessenberg	$2n^3/3 + O(n^2)$[10]:474	$4n^3/3 + O(n^2)$[10]:474	Reflect each column through a subspace to zero out its lower entries.
Givens rotations	General	Hessenberg	$4n^3/3 + O(n^2)$[10]:470		Apply planar rotations to zero out individual entries. Rotations are ordered so that later ones do not cause zero entries to become non-zero again.
Arnoldi iteration	General	Hessenberg			Perform Gram–Schmidt orthogonalization on Krylov subspaces.
Lanczos algorithm	Hermitian	Tridiagonal			Arnoldi iteration for Hermitian matrices, with shortcuts.

For symmetric tridiagonal eigenvalue problems all eigenvalues (without eigenvectors) can be computed numerically in time O(n log(n)), using bisection on the characteristic polynomial.[11]

Iterative algorithms

Iterative algorithms solve the eigenvalue problem by producing sequences that converge to the eigenvalues. Some algorithms also produce sequences of vectors that converge to the eigenvectors. Most commonly, the eigenvalue sequences are expressed as sequences of similar matrices which converge to a triangular or diagonal form, allowing the eigenvalues to be read easily. The eigenvector sequences are expressed as the corresponding similarity matrices.

Method	Applies to	Produces	Cost per step	Convergence	Description
Lanczos algorithm	Hermitian	m largest/smallest eigenpairs			
Power iteration	general	eigenpair with largest value	$O(n^2)$	linear	Repeatedly applies the matrix to an arbitrary starting vector and renormalizes.
Inverse iteration	general	eigenpair with value closest to μ		linear	Power iteration for $(A - \mu I)^{-1}$
Rayleigh quotient iteration	Hermitian	any eigenpair		cubic	Power iteration for $(A - \mu_i I)^{-1}$, where μ_i for each iteration is the Rayleigh quotient of the previous iteration.
Preconditioned inverse iteration[12] or LOBPCG algorithm	positive-definite real symmetric	eigenpair with value closest to μ			Inverse iteration using a preconditioner (an approximate inverse to A).
Bisection method	real symmetric tridiagonal	any eigenvalue		linear	Uses the bisection method to find roots of the characteristic polynomial, supported by the Sturm sequence.
Laguerre iteration	real symmetric tridiagonal	any eigenvalue		cubic[13]	Uses Laguerre's method to find roots of the characteristic polynomial, supported by the Sturm sequence.
QR algorithm	Hessenberg	all eigenvalues	$O(n^2)$	cubic	Factors $A = QR$, where Q is orthogonal and R is triangular, then applies the next iteration to RQ.
		all eigenpairs	$6n^3 + O(n^2)$		
Jacobi eigenvalue algorithm	real symmetric	all eigenvalues	$O(n^3)$	quadratic	Uses Givens rotations to attempt clearing all off-diagonal entries. This fails, but strengthens the diagonal.
Divide-and-conquer	Hermitian tridiagonal	all eigenvalues	$O(n^2)$		Divides the matrix into submatrices that are diagonalized

		all eigenpairs	$(4/3)n^3 + O(n^2)$		then recombined.
Homotopy method	real symmetric tridiagonal	all eigenpairs	$O(n^2)$[14]		Constructs a computable homotopy path from a diagonal eigenvalue problem.
Folded spectrum method	real symmetric	eigenpair with value closest to μ			Preconditioned inverse iteration applied to $(A - \mu I)^2$
MRRR algorithm[15]	real symmetric tridiagonal	some or all eigenpairs	$O(n^2)$		"Multiple relatively robust representations" – performs inverse iteration on a LDL^T decomposition of the shifted matrix.
Gram iteration[16]	general	Eigenpair with largest eigenvalue		super-linear	Repeatedly computes the Gram product and rescales, deterministically.

Direct calculation

While there is no simple algorithm to directly calculate eigenvalues for general matrices, there are numerous special classes of matrices where eigenvalues can be directly calculated. These include:

Triangular matrices

Since the determinant of a triangular matrix is the product of its diagonal entries, if T is triangular, then $\det(\lambda I - T) = \prod_i (\lambda - T_{ii})$. Thus the eigenvalues of T are its diagonal entries.

Factorable polynomial equations

If p is any polynomial and $p(A) = 0$, then the eigenvalues of A also satisfy the same equation. If p happens to have a known factorization, then the eigenvalues of A lie among its roots.

For example, a projection is a square matrix P satisfying $P^2 = P$. The roots of the corresponding scalar polynomial equation, $\lambda^2 = \lambda$, are 0 and 1. Thus any projection has 0 and 1 for its eigenvalues. The multiplicity of 0 as an eigenvalue is the nullity of P, while the multiplicity of 1 is the rank of P.

Another example is a matrix A that satisfies $A^2 = \alpha^2 I$ for some scalar α. The eigenvalues must be $\pm\alpha$. The projection operators

$$P_+ = \frac{1}{2}\left(I + \frac{A}{\alpha}\right)$$

$$P_- = \frac{1}{2}\left(I - \frac{A}{\alpha}\right)$$

satisfy

$$AP_+ = \alpha P_+ \quad AP_- = -\alpha P_-$$

and

$$P_+P_+ = P_+ \quad P_-P_- = P_- \quad P_+P_- = P_-P_+ = 0.$$

The column spaces of P_+ and P_- are the eigenspaces of A corresponding to $+\alpha$ and $-\alpha$, respectively.

2×2 matrices

For dimensions 2 through 4, formulas involving radicals exist that can be used to find the eigenvalues. While a common practice for 2×2 and 3×3 matrices, for 4×4 matrices the increasing complexity of the root formulas makes this approach less attractive.

For the 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

the characteristic polynomial is

$$\det\begin{bmatrix} \lambda - a & -b \\ -c & \lambda - d \end{bmatrix} = \lambda^2 - (a+d)\,\lambda + (ad - bc) = \lambda^2 - \lambda\,\mathrm{tr}(A) + \det(A).$$

Thus the eigenvalues can be found by using the quadratic formula:

$$\lambda = \frac{\mathrm{tr}(A) \pm \sqrt{\mathrm{tr}^2(A) - 4\det(A)}}{2}.$$

Defining $\mathrm{gap}\,(A) = \sqrt{\mathrm{tr}^2(A) - 4\det(A)}$ to be the distance between the two eigenvalues, it is straightforward to calculate

$$\frac{\partial\lambda}{\partial a} = \frac{1}{2}\left(1 \pm \frac{a-d}{\mathrm{gap}(A)}\right), \qquad \frac{\partial\lambda}{\partial b} = \frac{\pm c}{\mathrm{gap}(A)}$$

with similar formulas for c and d. From this it follows that the calculation is well-conditioned if the eigenvalues are isolated.

Eigenvectors can be found by exploiting the Cayley–Hamilton theorem. If λ_1, λ_2 are the eigenvalues, then $(A - \lambda_1 I)(A - \lambda_2 I) = (A - \lambda_2 I)(A - \lambda_1 I) = 0$, so the columns of $(A - \lambda_2 I)$ are annihilated by $(A - \lambda_1 I)$ and vice versa. Assuming neither matrix is zero, the columns of each must include eigenvectors for the other eigenvalue. (If either matrix is zero, then A is a multiple of the identity and any non-zero vector is an eigenvector.)

For example, suppose

$$A = \begin{bmatrix} 4 & 3 \\ -2 & -3 \end{bmatrix},$$

then $\operatorname{tr}(A) = 4 - 3 = 1$ and $\det(A) = 4(-3) - 3(-2) = -6$, so the characteristic equation is

$$0 = \lambda^2 - \lambda - 6 = (\lambda - 3)(\lambda + 2),$$

and the eigenvalues are 3 and -2. Now,

$$A - 3I = \begin{bmatrix} 1 & 3 \\ -2 & -6 \end{bmatrix}, \qquad A + 2I = \begin{bmatrix} 6 & 3 \\ -2 & -1 \end{bmatrix}.$$

In both matrices, the columns are multiples of each other, so either column can be used. Thus, $(1, -2)$ can be taken as an eigenvector associated with the eigenvalue -2, and $(3, -1)$ as an eigenvector associated with the eigenvalue 3, as can be verified by multiplying them by A.

3×3 matrices

The characteristic equation of a symmetric 3×3 matrix A is:

$$\det(\alpha I - A) = \alpha^3 - \alpha^2 \operatorname{tr}(A) - \alpha \frac{1}{2}\left(\operatorname{tr}(A^2) - \operatorname{tr}^2(A)\right) - \det(A) = 0.$$

This equation may be solved using the methods of Cardano or Lagrange, but an affine change to A will simplify the expression considerably, and lead directly to a trigonometric solution. If $A = pB + qI$, then A and B have the same eigenvectors, and β is an eigenvalue of B if and only if $\alpha = p\beta + q$ is an eigenvalue of A. Letting $q = \operatorname{tr}(A)/3$ and $p = \left(\operatorname{tr}\left((A - qI)^2\right)/6\right)^{1/2}$, gives

$$\det(\beta I - B) = \beta^3 - 3\beta - \det(B) = 0.$$

The substitution $\beta = 2\cos\theta$ and some simplification using the identity $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$ reduces the equation to $\cos 3\theta = \det(B) / 2$. Thus

$$\beta = 2\cos\left(\frac{1}{3}\arccos\left(\det(B)/2\right) + \frac{2k\pi}{3}\right), \quad k = 0, 1, 2.$$

If $\det(B)$ is complex or is greater than 2 in absolute value, the arccosine should be taken along the same branch for all three values of k. This issue doesn't arise when A is real and symmetric, resulting in a simple algorithm:[17]

```
% Given a real symmetric 3x3 matrix A, compute the eigenvalues
% Note that acos and cos operate on angles in radians

p1 = A(1,2)^2 + A(1,3)^2 + A(2,3)^2
if (p1 == 0)
   % A is diagonal.
   eig1 = A(1,1)
   eig2 = A(2,2)
   eig3 = A(3,3)
else
   q = trace(A)/3              % trace(A) is the sum of all diagonal values
   p2 = (A(1,1) - q)^2 + (A(2,2) - q)^2 + (A(3,3) - q)^2 + 2 * p1
   p = sqrt(p2 / 6)
   B = (1 / p) * (A - q * I)    % I is the identity matrix
   r = det(B) / 2

   % In exact arithmetic for a symmetric matrix  -1 <= r <= 1
   % but computation error can leave it slightly outside this range.
```

```
    if (r <= -1)
       phi = pi / 3
    elseif (r >= 1)
       phi = 0
    else
       phi = acos(r) / 3
    end

    % the eigenvalues satisfy eig3 <= eig2 <= eig1
    eig1 = q + 2 * p * cos(phi)
    eig3 = q + 2 * p * cos(phi + (2*pi/3))
    eig2 = 3 * q - eig1 - eig3     % since trace(A) = eig1 + eig2 + eig3
end
```

Once again, the eigenvectors of A can be obtained by recourse to the Cayley–Hamilton theorem. If α_1, α_2, α_3 are distinct eigenvalues of A, then $(A - \alpha_1 I)(A - \alpha_2 I)(A - \alpha_3 I) = 0$. Thus the columns of the product of any two of these matrices will contain an eigenvector for the third eigenvalue. However, if $\alpha_3 = \alpha_1$, then $(A - \alpha_1 I)^2(A - \alpha_2 I) = 0$ and $(A - \alpha_2 I)(A - \alpha_1 I)^2 = 0$. Thus the *generalized* eigenspace of α_1 is spanned by the columns of $A - \alpha_2 I$ while the ordinary eigenspace is spanned by the columns of $(A - \alpha_1 I)(A - \alpha_2 I)$. The ordinary eigenspace of α_2 is spanned by the columns of $(A - \alpha_1 I)^2$.

For example, let

$$A = \begin{bmatrix} 3 & 2 & 6 \\ 2 & 2 & 5 \\ -2 & -1 & -4 \end{bmatrix}.$$

The characteristic equation is

$$0 = \lambda^3 - \lambda^2 - \lambda + 1 = (\lambda - 1)^2(\lambda + 1),$$

with eigenvalues 1 (of multiplicity 2) and -1. Calculating,

$$A - I = \begin{bmatrix} 2 & 2 & 6 \\ 2 & 1 & 5 \\ -2 & -1 & -5 \end{bmatrix}, \qquad A + I = \begin{bmatrix} 4 & 2 & 6 \\ 2 & 3 & 5 \\ -2 & -1 & -3 \end{bmatrix}$$

and

$$(A - I)^2 = \begin{bmatrix} -4 & 0 & -8 \\ -4 & 0 & -8 \\ 4 & 0 & 8 \end{bmatrix}, \qquad (A - I)(A + I) = \begin{bmatrix} 0 & 4 & 4 \\ 0 & 2 & 2 \\ 0 & -2 & -2 \end{bmatrix}$$

Thus $(-4, -4, 4)$ is an eigenvector for -1, and $(4, 2, -2)$ is an eigenvector for 1. $(2, 3, -1)$ and $(6, 5, -3)$ are both generalized eigenvectors associated with 1, either one of which could be combined with $(-4, -4, 4)$ and $(4, 2, -2)$ to form a basis of generalized eigenvectors of A. Once found, the eigenvectors can be normalized if needed.

Eigenvectors of normal 3×3 matrices

If a 3×3 matrix $\boldsymbol{A}$ is normal, then the cross-product can be used to find eigenvectors. If $\boldsymbol{\lambda}$ is an eigenvalue of $\boldsymbol{A}$, then the null space of $\boldsymbol{A - \lambda I}$ is perpendicular to its column space. The cross product of two independent columns of $\boldsymbol{A - \lambda I}$ will be in the null space. That is, it will be an eigenvector associated with $\boldsymbol{\lambda}$. Since the column space is two dimensional in this case, the eigenspace must be one dimensional, so any other eigenvector will be parallel to it.

If $A - \lambda I$ does not contain two independent columns but is not $\mathbf{0}$, the cross-product can still be used. In this case λ is an eigenvalue of multiplicity 2, so any vector perpendicular to the column space will be an eigenvector. Suppose $\mathbf{v}$ is a non-zero column of $A - \lambda I$. Choose an arbitrary vector $\mathbf{u}$ not parallel to $\mathbf{v}$. Then $\mathbf{v} \times \mathbf{u}$ and $(\mathbf{v} \times \mathbf{u}) \times \mathbf{v}$ will be perpendicular to $\mathbf{v}$ and thus will be eigenvectors of λ.

This does not work when A is not normal, as the null space and column space do not need to be perpendicular for such matrices.

See also

- List of eigenvalue algorithms

Notes

1. The term "ordinary" is used here only to emphasize the distinction between "eigenvector" and "generalized eigenvector".
2. where the constant term is multiplied by the identity matrix I.
3. This ordering of the inner product (with the conjugate-linear position on the left), is preferred by physicists. Algebraists often place the conjugate-linear position on the right: $\mathbf{w} \cdot \mathbf{v} = \mathbf{v}^* \mathbf{w}$.

References

1. Axler, Sheldon (1995), "Down with Determinants!" (https://web.archive.org/web/20120913111605/http://www.axler.net/DwD.pdf) (PDF), *American Mathematical Monthly*, **102** (2): 139–154, doi:10.2307/2975348 (https://doi.org/10.2307%2F2975348), JSTOR 2975348 (https://www.jstor.org/stable/2975348), archived from the original (http://www.axler.net/DwD.pdf) (PDF) on 2012-09-13, retrieved 2012-07-31
2. F. L. Bauer; C. T. Fike (1960), "Norms and exclusion theorems", *Numer. Math.*, **2**: 137–141, doi:10.1007/bf01386217 (https://doi.org/10.1007%2Fbf01386217), S2CID 121278235 (https://api.semanticscholar.org/CorpusID:121278235)
3. S.C. Eisenstat; I.C.F. Ipsen (1998), "Relative Perturbation Results for Eigenvalues and Eigenvectors of Diagonalisable Matrices" (http://www.lib.ncsu.edu/resolver/1840.4/286), *BIT*, **38** (3): 502–9, doi:10.1007/bf02510256 (https://doi.org/10.1007%2Fbf02510256), S2CID 119886389 (https://api.semanticscholar.org/CorpusID:119886389)
4. J. Dongarra and F. Sullivan (2000). "Top ten algorithms of the century". *Computing in Science and Engineering*. **2**: 22-23. doi:10.1109/MCISE.2000.814652 (https://doi.org/10.1109%2FMCISE.2000.814652).
5. Thompson, R. C. (June 1966). "Principal submatrices of normal and Hermitian matrices" (https://doi.org/10.1215%2Fijm%2F1256055111). *Illinois Journal of Mathematics*. **10** (2): 296–308. doi:10.1215/ijm/1256055111 (https://doi.org/10.1215%2Fijm%2F1256055111).
6. Peter Nylen; Tin-Yau Tam; Frank Uhlig (1993). "On the eigenvalues of principal submatrices of normal, hermitian and symmetric matrices". *Linear and Multilinear Algebra*. **36** (1): 69–78. doi:10.1080/03081089308818276 (https://doi.org/10.1080%2F03081089308818276).
7. Bebiano N, Furtado S, da Providência J (2011). "On the eigenvalues of principal submatrices of J-normal matrices" (https://doi.org/10.1016%2Fj.laa.2011.05.033). *Linear Algebra and Its Applications*. **435** (12): 3101–3114. doi:10.1016/j.laa.2011.05.033 (https://doi.org/10.1016%2Fj.laa.2011.05.033).
8. Forrester PJ, Zhang J (2021). "Corank-1 projections and the randomised Horn problem". *Tunisian Journal of Mathematics*. **3**: 55–73. arXiv:1905.05314 (https://arxiv.org/abs/1905.05314). doi:10.2140/tunis.2021.3.55 (https://doi.org/10.2140%2Ftunis.2021.3.55). S2CID 153312446 (https://api.semanticscholar.org/CorpusID:153312446).

9. Denton PB, Parke SJ, Tao T, Zhang X (2021). "Eigenvectors from eigenvalues: A survey of a basic identity in linear algebra". *Bulletin of the American Mathematical Society*. **59**: 1. arXiv:1908.03795 (https://arxiv.org/abs/1908.03795). doi:10.1090/bull/1722 (https://doi.org/10.1090%2Fbull%2F1722). S2CID 213918682 (https://api.semanticscholar.org/CorpusID:213918682).
10. Press, William H.; Teukolsky, Saul A.; Vetterling, William T.; Flannery, Brian P. (1992). *Numerical Recipes in C* (https://archive.org/details/numericalrecipes00pres_0) (2nd ed.). Cambridge University Press. ISBN 978-0-521-43108-8.
11. Coakley, Ed S. (May 2013), "A fast divide-and-conquer algorithm for computing the spectra of real symmetric tridiagonal matrices.", *Applied and Computational Harmonic Analysis*, **34** (3): 379–414, doi:10.1016/j.acha.2012.06.003 (https://doi.org/10.1016%2Fj.acha.2012.06.003)
12. Neymeyr, K. (2006), "A geometric theory for preconditioned inverse iteration IV: On the fastest convergence cases.", *Linear Algebra Appl.*, **415** (1): 114–139, doi:10.1016/j.laa.2005.06.022 (https://doi.org/10.1016%2Fj.laa.2005.06.022)
13. Li, T. Y.; Zeng, Zhonggang (1992), "Laguerre's Iteration In Solving The Symmetric Tridiagonal Eigenproblem - Revisited", *SIAM Journal on Scientific Computing*
14. Chu, Moody T. (1988), "A Note on the Homotopy Method for Linear Algebraic Eigenvalue Problems", *Linear Algebra Appl.*, **105**: 225–236, doi:10.1016/0024-3795(88)90015-8 (https://doi.org/10.1016%2F0024-3795%2888%2990015-8)
15. Dhillon, Inderjit S.; Parlett, Beresford N.; Vömel, Christof (2006), "The Design and Implementation of the MRRR Algorithm" (http://www.cs.utexas.edu/users/inderjit/public_papers/DesignMRRR_toms06.pdf) (PDF), *ACM Transactions on Mathematical Software*, **32** (4): 533–560, doi:10.1145/1186785.1186788 (https://doi.org/10.1145%2F1186785.1186788), S2CID 2410736 (https://api.semanticscholar.org/CorpusID:2410736)
16. Delattre, B.; Barthélemy, Q.; Araujo, A.; Allauzen, A. (2023), "Efficient Bound of Lipschitz Constant for Convolutional Layers by Gram Iteration" (https://proceedings.mlr.press/v202/delattre23a.html), *Proceedings of the 40th International Conference on Machine Learning*
17. Smith, Oliver K. (April 1961), "Eigenvalues of a symmetric 3 × 3 matrix.", *Communications of the ACM*, **4** (4): 168, doi:10.1145/355578.366316 (https://doi.org/10.1145%2F355578.366316), S2CID 37815415 (https://api.semanticscholar.org/CorpusID:37815415)

Further reading

- Bojanczyk, Adam W.; Adam Lutoborski (Jan 1991). "Computation of the Euler angles of a symmetric 3X3 matrix" (http://cacm.acm.org/magazines/1961/4/14532-eigenvalues-of-a-symmetric-3-%C3%83-3-matrix/abstract). *SIAM Journal on Matrix Analysis and Applications*. **12** (1): 41–48. doi:10.1137/0612005 (https://doi.org/10.1137%2F0612005).
