A.6 Computing eigenvalues

As discussed above, eigenvalues are the characteristic values that ensure $(A - \lambda I)$ has a nullspace for square matrix A. That is, $(A - \lambda I)\,x = 0$ where x is an eigenvector. If an eigenvector can be identified such that $Ax = \lambda x$ then the constant, λ, is an associated eigenvalue. For instance, if the rows of A have the same sum then $x = \iota$ (a vector of ones) and λ equals the sum of any row of A.

Further, since the sum of the eigenvalues equals the trace of the matrix and the product of the eigenvalues equals the determinant of the matrix, finding the eigenvalues for small matrices is relatively simple. For instance, eigenvalues of a 2×2 matrix can be found by solving

$$\begin{aligned} \lambda_1 + \lambda_2 &= tr\,(A) \\ \lambda_1 \lambda_2 &= \det(A) \end{aligned}$$

Alternatively, we can solve the roots or zeroes of the characteristic polynomial. That is, $\det(A - \lambda I) = 0$.

Example 1 *Suppose* $A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$ *then* $tr\,(A) = 5$ *and* $\det(A) = 4$. *Therefore,*

$$\begin{aligned} \lambda_1 + \lambda_2 &= 5 \\ \lambda_1 \lambda_2 &= 4 \end{aligned}$$

which leads to $\lambda_1 = 4$ *and* $\lambda_2 = 1$. *Likewise, the characteristic polynomial is* $\det(A - \lambda I) = (2 - \lambda)(3 - \lambda) - 2 = 0$ *leading to the same solution for* λ.

However, for larger matrices this approach proves impractical. Hence, we'll explore some alternatives.

A.6.1 Schur's lemma

Schur's lemma says that while every square matrix may not be diagonalizable, it can be triangularized by some unitary operator U.

$$\begin{aligned} T &= U^{-1} A U \\ &= U^{*} A U \end{aligned}$$

or

$$A = U T U^{*}$$

where A is the matrix of interest, T is a triangular matrix, and U is unitary so that $U^{*}U = UU^{*} = I$ (U^{*} denotes the complex conjugate transpose of

U). Further, since T and A are similar matrices they have the same eigenvalues and the eigenvalues reside on the main diagonal of T. To see they are similar matrices recognize they have the same characteristic polynomial.

$$\begin{aligned}
\det(A-\lambda I) &= \det(T-\lambda I) \\
&= \det(U^*AU-\lambda I) \\
&= \det(U^*AU-\lambda U^*IU) \\
&= \det(U^*(A-\lambda I)U) \\
&= \det(U^*)\det(A-\lambda I)\det(U) \\
&= 1\det(A-\lambda I)\,1 \\
&= \det(A-\lambda I)
\end{aligned}$$

Before discussing construction of T, we introduce some eigenvalue construction algorithms.

A.6.2 Power algorithm

The power algorithm is an iterative process for finding the largest absolute value eigenvalue.

1. Let k_1 be a vector of ones where the number of elements in the vector equals the number of rows or columns in A.
2. Let $k_{t+1} = \frac{Ak_t}{\sqrt{k_t^T A^T A k_t}}$ where $\sqrt{k_t^T A^T A k_t} = norm$.
3. iterate until $|k_{t+1} - k_t| < \varepsilon\iota$ for desired precision ε.
4. $norm$ is the largest eigenvalue of A and $k_t = k_{t+1}$ is it's associated eigenvector.

Clearly, if $k_t = k_{t+1}$ this satisfies the property of eigenvalues and eigenvectors, $Ax = \lambda x$ or $Ak_t = \sqrt{k_t^T A^T A k_t}\, k_t$.

Example 2 *Continue with* $A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$. $k_2 = \frac{Ak_1}{norm_1} = \frac{1}{4\sqrt{2}} \begin{bmatrix} 4 \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ $k_3 = \frac{Ak_2}{norm_2} = \frac{1}{4} \begin{bmatrix} \frac{4}{\sqrt{2}} \\ \frac{4}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ *Hence,* $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ *is an eigenvector and* $norm_2 = 4$ *is the associated (largest) eigenvalue.*

Example 3 (complex eigenvalues) *Suppose* $A = \begin{bmatrix} -4 & 2 \\ -2 & -4 \end{bmatrix}$. *The eigenvalues are* $\lambda = -4 \pm 2i$ *with* $norm = \sqrt{(-4+2i)(-4-2i)} = 4.472136$ *(not a complex number). The power algorithm settles on the norm but* $Ak_n \neq norm * k_n$. *Try the algorithm again except begin with* $k_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$. *The algorithm converges to the same norm but* $k_n = \begin{bmatrix} -0.4406927 - 0.5529828i \\ 0.5529828 - 0.4406927i \end{bmatrix}$.

Now,

$$\begin{aligned} Ak_n &= \lambda k_n \\ & \begin{bmatrix} -4 & 2 \\ -2 & -4 \end{bmatrix} \begin{bmatrix} -0.4406927 - 0.5529828i \\ 0.5529828 - 0.4406927i \end{bmatrix} \\ &= \lambda \begin{bmatrix} -0.4406927 - 0.5529828i \\ 0.5529828 - 0.4406927i \end{bmatrix} \end{aligned}$$

solving for λ yields $-4 + 2i$. Since complex roots always come in conjugate pairs we also know the other eigenvalue, $-4 - 2i$.

A.6.3 QR algorithm

The QR algorithm parallels Schur's lemma and supplies a method to compute all eigenvalues.

1. Compute the factors Q, an orthogonal matrix $QQ^T = Q^T Q = I$, and R, a right or upper triangular matrix, such that $A = QR$.
2. Reverse the factors and denote this A_1, $A_1 = RQ$.
3. Factor A_1, $A_1 = Q_1 R_1$ then $A_2 = R_1 Q_1$.
4. Repeat until A_k is triangular.

$$\begin{aligned} A_{k-1} &= Q_{k-1} R_{k-1} \\ A_k &= R_{k-1} Q_{k-1} \end{aligned}$$

The main diagonal elements of A_k are the eigenvalues of A.

The connection to Schur's lemma is $RQ = Q^T QRQ = Q^T AQ = A_1$ so that A, A_1 and A_k are similar matrices (they have the same eigenvalues).

Example 4 *Continue with* $A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$. $A_1 = RQ = \begin{bmatrix} 3.4 & -1.8 \\ -0.8 & 1.6 \end{bmatrix}$ *and* $A_{11} = R_{10} Q_{10} = \begin{bmatrix} 4 & -1 \\ 0 & 1 \end{bmatrix}$.[17] *Hence, the eigenvalues of A (and also A_{10}) are the main diagonal elements, 4 and 1.*

Example 5 (complex eigenvalues) *Suppose* $A = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 3 \\ 0 & -3 & 2 \end{bmatrix}$. *The QR algorithm leaves A unchanged. However, we can work in blocks to solve for the eigenvalues. The first block is simply bordered by zeroes or the first row, first column element ($B_1 = 5$) and 5 is an eigenvalue. The second block is rows 2 and 3 and columns 2 and 3 or* $B_2 = \begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix}$. *Now solve*

[17] Shifting refinements are typically employed to speed convergence (see Strang).

the characteristic polynomial for this 2×2 matrix.

$$\begin{aligned} -\lambda^2 + 4\lambda - 13 &= 0 \\ \lambda &= 2 \pm 3i \end{aligned}$$

We can check that each of these three eigenvalues creates a nullspace for $A - \lambda I$.

$$A - 5I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -3 & 3 \\ 0 & -3 & -3 \end{bmatrix}$$

has rank 2 and nullspace or eigenvector $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

$$A - (2+3i)\, I = \begin{bmatrix} 3-3i & 0 & 0 \\ 0 & -3i & 3 \\ 0 & -3 & -3i \end{bmatrix}$$

The second row is a scalar multiple $(-i)$ of the third (and vice versa) and a nullspace or eigenvector is $\frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ i \\ -1 \end{bmatrix}$. *Finally,*[18]

$$A - (2-3i)\, I = \begin{bmatrix} 3-3i & 0 & 0 \\ 0 & 3i & 3 \\ 0 & -3 & 3i \end{bmatrix}$$

Again, the second row is a scalar multiple (i) of the third (and vice versa) and a nullspace or eigenvector is $\frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ i \\ 1 \end{bmatrix}$. *Hence, the eigenvalues are $\lambda = 5, 2 \pm 3i$.*

A.6.4 Schur decomposition

Schur decomposition works similarly.

1. Use one of the above algorithms to find an eigenvalue of $n \times n$ matrix A, λ_1.
2. From this eigenvalue, construct a unit length eigenvector, x_1.

[18] Gauss' fundamental theorem of algebra insures complex roots always come in conjugate pairs so this may be overly pedantic.

3. Utilize Gram-Schmidt to construct a unitary matrix U_1 from $n-1$ columns of A where x_1 is the first column of U. This creates

$$AU_1 = U_1 \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

or

$$U_1^* A U_1 = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

4. The next step works the same way except with the lower right $(n-1) \times (n-1)$ matrix. then, U_2 is constructed from this lower, right block with a one in the upper, left position with zeroes in its row and column.

$$U_2 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & x_{22} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & x_{2n} & \cdots & * \end{bmatrix}$$

$$U_2^* U_1^* A U_1 U_2 = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & * \end{bmatrix}$$

5. Continue until T is constructed.

$$\begin{aligned} T &= U_{n-1}^* \cdots U_1^* A U_1 \cdots U_{n-1} \\ U^* A U &= \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \end{aligned}$$

where $U = U_1 \cdots U_{n-1}$.

Example 6 (not diagonalizable) *Suppose* $A = \begin{bmatrix} 5 & 0 & 1 \\ 0 & 2 & -3 \\ 0 & -3 & 2 \end{bmatrix}$. *This matrix has repeated eigenvalues* $(5, 5, -1)$ *and lacks a full set of linearly indepedent eigenvectors therefore it cannot be expressed in diagonalizable form* $A = S\Lambda S^{-1}$ *(as the latter term doesn't exist). Nonetheless, the Schur decomposition can still be employed to triangularize the matrix. A unit length* *eigenvector associated with* $\lambda = 5$ *is* $x_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. *Applying Gram-Schmidt to columns two and three of A yields* $U_1 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0.55470 & -0.83205 \\ 0 & -0.83205 & -0.55470 \end{bmatrix}$. *This leads to*

$$\begin{aligned} T_1 &= U_1^* A U_1 \\ &= \begin{bmatrix} 5 & -0.83205 & -0.55470 \\ 0 & 4.76923 & -1.15385 \\ 0 & -1.15385 & -0.76923 \end{bmatrix} \end{aligned}$$

Working with the lower, right 2×2 *block gives*

$$U_2 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -0.98058 & -0.19612 \\ 0 & 0.19612 & -0.98058 \end{bmatrix}$$

Then,

$$\begin{aligned} T &= U_2^* U_1^* A U_1 U_2 \\ U^* A U &= \begin{bmatrix} 5 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 5 & 0 \\ 0 & 0 & -1 \end{bmatrix} \end{aligned}$$

where $U = U_1 U_2 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$.

Example 7 (complex eigenvalues) *Suppose* $A = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 3 \\ 0 & -3 & 2 \end{bmatrix}$. *We know from example 5 A has complex eigenvalues. Let's explore its Schur decomposition. Again,* $\lambda = 5$ *is an eigenvalue with corresponding eigenvector* $x_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. *Applying Gram-Schmidt to columns two and three of A yields* $U_1 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0.55470 & 0.83205 \\ 0 & -0.83205 & 0.55470 \end{bmatrix}$. *This leads to*

$$\begin{aligned} T_1 &= U_1^* A U_1 \\ &= \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 3 \\ 0 & -3 & 2 \end{bmatrix} \end{aligned}$$

Working with the lower, right 2×2 *block,* $\lambda = 2 + 3i$*, and associated eigenvector* $x_2 = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}}i \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$ *gives*

$$U_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}}i & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}i \end{bmatrix}$$

where $x_{12} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{2}}i \\ 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}$ *is applied via Gram-Schmidt to create the third (column) vector of* U_2 *from the third column of* A*,* $A_{\cdot 3}$*.*[19]

$$\begin{aligned} & A_{\cdot 3} - x_{12}x_{12}^* A_{\cdot 3} \\ = & \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{2}}i \\ 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{\sqrt{2}}i & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ -3i \end{bmatrix} \end{aligned}$$

before normalization and after we have $\begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}}i \end{bmatrix}$*. Then,*

$$\begin{aligned} T &= U_2^* U_1^* A U_1 U_2 \\ U^* A U &= \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2+3i & 0 \\ 0 & 0 & 2-3i \end{bmatrix} \end{aligned}$$

where $U = U_1 U_2 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -0.5883484 + 0.3922323i & 0.3922323 - 0.5883484i \\ 0 & -0.3922323 - 0.5883484i & -0.5883484 - 0.3922323i \end{bmatrix}$.
The eigenvalues lie along the main diagonal of T*.*

[19] Notice, conjugate transpose is employed in the construction of the projection matrix to accommodate complex elements.