1 Similar Matrices & QR Iteration Introduction (5.4)

Two matrices, $A, B \in \mathbb{R}^{n\times n}$, are **similar** if there exists an invertible matrix $S \in \mathbb{R}^{n\times n}$ such that $AS = SB$. Equivalently,

$$A = SBS^{-1} \qquad B = S^{-1}AS.$$

A and B are called **orthogonally similar** if S is orthogonal and $A = SBS^{-1}$. In this case, we actually have $A = SBS^T$.

Theorem 1.1

Similar matrices have the same eigenvalues.

That is, if $B = S^{-1}AS$ and v is an eigenvector of A to the eigenvalue λ, then $S^{-1}v$ is an eigenvector of B with respect to λ.

Proof. We have

$$\begin{aligned} Av = \lambda v &\implies SBS^{-1}v = \lambda v \\ &\implies S^{-1}(SBS^{-1})v = S^{-1}(\lambda v) \\ &\implies (S^{-1}S)BS^{-1}v = \lambda S^{-1}v \\ &\implies BS^{-1}v = \lambda S^{-1}v. \end{aligned} \tag{1}$$

This means that $S^{-1}v$ is an eigenvalue of B with respect to λ. □

Lemma 1.1: Diagonalizing Semisimple Matrices

Let A be a matrix in $\mathbb{R}^{n\times n}$. A is semisimple if and only if there exists an invertible matrix $V \in \mathbb{R}^{n\times n}$ and a diagonal matrix $D \in \mathbb{R}^{n\times n}$ such that

$$A = VDV^{-1}.$$

Remarks:

- A and D are similar, meaning they have the same eigenvalues[1].
- $A = VDV^{-1}$ is equivalent to $AV = VD$. This is an eigenvalue/vector equation.

Proof Idea. If $D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$ contains eigenvalues and $V = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$ contains eigenvectors, then V is invertible implies that $\begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$ are linearly independent. Thus, A is semisimple. How do we find the eigenvalues of A?

1.1 Interlude: Complex Matrices

Let $\alpha = a + bi \in \mathbb{C}$, where $a, b \in \mathbb{R}$. Then,

- The complex conjugate, $\bar{\alpha} = a - bi \in \mathbb{C}$.
- Also, $|\alpha| = \sqrt{a^2 + b^2}$.

Regarding complex matrices,

- **Generalization of Transpose:** Let $A \in \mathbb{C}^{n\times n}$. Then, $A^* = \bar{A}^T$, known as the generalization of a transposition. $\bar{A}$ is the complex conjugate of every entry.

[1] The eigenvalues of a diagonal matrix is just the entries on the diagonal.

- **Generalization of Orthogonality:** Recall that $Q \in \mathbb{R}^{n\times n}$ is orthogonal if $QQ^T = Q^TQ = I$. How do we generalize this to complex matrices? Let $U \in \mathbb{C}^{n\times n}$. U is called **unitary** if $UU^* = U^*U = I$.

Theorem 1.2: Schur

Let $A \in \mathbb{C}^{n\times n}$. Then, there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ and an upper triangular matrix $T \in \mathbb{C}^{n\times n}$ such that $A = UTU^*$.

Remark: A is unitarily similar to T. So, A and T has the same eignevalues.

1.2 Back to Real Matrices

If $A \in \mathbb{R}^{n\times n}$, we can still apply Schur's theorem[2]; that is,

$$A = UTU^* \qquad U, T \in \mathbb{C}^{n\times n}.$$

Another version of Schur's Theorem, known as the Real Schur's Theorem, states the following.

Theorem 1.3: Real Schur

If $A \in \mathbb{R}^{n\times n}$. Then, there exists an orthogonal $Q \in \mathbb{R}^{n\times n}$ and an "almost" upper triangular matrix $T \in \mathbb{R}^{n\times n}$ such that $A = QTQ^T$.

We can think of diagonal entries of T as consisting of size 1×1 or size 2×2 blocks.

(Example.) If $A \in \mathbb{R}^{4\times 4}$ with eigenvalues $2+i$, $2-i$, 5, and 6. Then,

- the complex Schur is

$$T = \begin{bmatrix} 2+i & * & * & * \\ 0 & 2-i & * & * \\ 0 & 0 & 5 & * \\ 0 & 0 & 0 & 6 \end{bmatrix} \in \mathbb{C}^{4\times 4}.$$

  Note that $\alpha = a + bi$ and $\bar{\alpha} = a - bi$, so

$$\alpha\bar{\alpha} = \alpha^2 + b^2.$$

- the real Schur is

$$T = \begin{bmatrix} 2 & 1 & * & * \\ -1 & 2 & * & * \\ 0 & 0 & 5 & * \\ 0 & 0 & 0 & 6 \end{bmatrix} \in \mathbb{R}^{4\times 4}.$$

  Notice how we have the 2×2 "block" at the top-left corner, representing the complex eigenvalues $2+i$ and $2-i$, respectively.

Remark: Note that complex eigenvalues always come in **complex conjugate pairs**. If $a+bi$ is an complex eigenvalue, then $a-bi$ is also a complex eigenvalue.

1.3 QR Iteration: A Basic Idea

The aim is to find the eigenvalues of a matrix $A \in \mathbb{C}^{n\times n}$. The idea behind the iterative procedure is as follows:

1. Step-by-step transform A without changing eigenvalues (similarity transformation).
2. Change into an upper-triangular matrix (T from Schur).

This method is based on the QR decomposition that we discussed earlier in the quarter.

[2] Recall that $\mathbb{R} \subset \mathbb{C}$.

1.3.1 Basic Idea: The Reals

Consider[3] $A \in \mathbb{R}^{n\times n}$. We know that $A = QR$, where Q is an orthogonal matrix. Then,

$$A = QR \implies Q^T A = Q^T QR = R \implies Q^T AQ = RQ.$$

Here, A and RQ have the same eigenvalues.

So, the iterative procedure begins by defining $A_0 = A$. The new matrix in iteration is $A_1 = RQ$. A_0 and A_1 have the same eigenvalues, so we can continue the process. So, the iterative procedure can be described in detailed as follows:

1. Iteratively compute QR decomposition.
2. Change multiplication order.
3. This converges to T, the upper-triangular from Schur.

So, starting with A_1, and look for its eigenvalues. Let $A_0 = A$. We can define

$$\boxed{A_k = R_k Q_k},$$

with R_k, Q_k from the QR decomposition of A_{k-1}; in other words,

$$\boxed{A_{k-1} = Q_k R_k}.$$

Then, A_{k-1} and A_k have the same eigenvalues.

More formally,

1. Let $A_0 = A = Q_1R_1$. Then, $A_1 = R_1Q_1$. Here, A_0 and A_1 have the same eigenvalues.
2. Let $A_1 = Q_2R_2$. Then, $A_2 = R_2Q_2$. Here, A_1 and A_2 have the same eigenvalues and, in particular, A_0, A_1, A_2 all have the same eigenvalues.
3. Continue the process. . .

Eventually, $\lim_{k\mapsto\infty} A_k = T$, with T from the Schur decomposition.

At the end, the eigenvalues of A are on the diagonal of T. If A is real, then T is "almost" upper triangular (real Schur decomposition).

Because this is an iterative method, we need a stopping criterion[4].

1.3.2 Disadvantages

There are some significant disadvantages with doing QR iteration.

- **Flop Count:** QR decomposition needs $\mathcal{O}(n^3)$ flops[5], and we need one QR decomposition in every step of the iteration. This is too much work.
- **Convergence Rate:** Convergence may be slow if the eigenvalues are close together in the absolute value. In case of distinct eigenvalues, $|\lambda_1| > |\lambda_2| > \ldots > |\lambda_n|$, applying QR iteration to A means the elements below the diagonal goes to 0 by the following rate[6]

$$(a_{ij}^{(k)}) = \mathcal{O}\left(\left|\frac{\lambda_i}{\lambda_j}\right|^k\right)$$

  for $i > j$ (i below diagonal).

[3] The complex matrix works the same
[4] To be discussed later.
[5] n represents the size of the matrix.
[6] $a_{ij}^{(k)}$ means the entry (i, j) in matrix A_k.

1.4 Upper Hessenberg Matrix

Definition 1.1: Upper Hessenberg Matrix

n $n \times n$ matrix H is called **upper Hessenberg** if $h_{ij} = 0$ for $i > j + 1$.

Note: If H is also symmetric ($A = A^T$), then we get tridiagonal.