Finite-dimensional spectral theory

part I: from $\mathbb{C}^n$ to the Schur decomposition

Ed Bueler

MATH 617 Functional Analysis

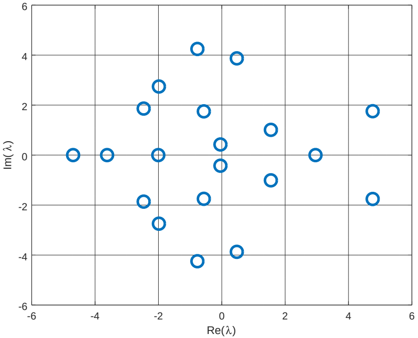
Spring 2020

linear algebra versus functional analysis

- these slides are about linear algebra, i.e. vector spaces of finite dimension, and linear operators on those spaces, i.e. matrices
- one definition of *functional analysis* might be: “rigorous extension of linear algebra to ∞-dimensional topological vector spaces”
    - it is important to understand the finite-dimensional case!
- the goal of these part I slides is to prove the Schur decomposition and the spectral theorem for matrices
- good references for these slides:
    - L. Trefethen & D. Bau, *Numerical Linear Algebra*, SIAM Press 1997
    - G. Strang, *Introduction to Linear Algebra*, 5th ed., Wellesley-Cambridge Press, 2016
    - G. Golub & C. van Loan, *Matrix Computations*, 4th ed., Johns Hopkins University Press 2013

- the *spectrum* $\sigma(A)$ of a square matrix A is its set of eigenvalues
    - reminder later about the definition of eigenvalues
    - $\sigma(A)$ is a subset of the complex plane $\mathbb{C}$
    - the plural of spectrum is *spectra*; the adjectival is *spectral*
- graphing $\sigma(A)$ gives the matrix a personality
    - example below: random, nonsymmetric, real 20×20 matrix

```
>> A = randn(20,20);
>> lam = eig(A);
>> plot(real(lam),imag(lam),'o')
>> grid on
>> xlabel('Re(\lambda)')
>> ylabel('Im(\lambda)')
```



$\mathbb{C}^n$ is an inner product space

- we use complex numbers $\mathbb{C}$ from now on
  - why? because eigenvalues can be complex even for a real matrix
  - recall: if $\alpha = x + iy \in \mathbb{C}$ then $\overline{\alpha} = x - iy$ is the *conjugate*
- let $\mathbb{C}^n$ be the space of (column) vectors with complex entries:

$$v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

- **definition.** an *inner product* on $\mathbb{C}^n$ is a function

$$\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C},$$

  almost-bilinear (*sesquilinear*[1]), with symmetry and positivity properties
- namely, for all $u, v, w \in \mathbb{C}^n$ and $\alpha \in \mathbb{C}$,
  - $\langle w, u + v\rangle = \langle w, u\rangle + \langle w, v\rangle$
  - $\langle u, \alpha v\rangle = \alpha \langle u, v\rangle$
  - $\langle u, v\rangle = \overline{\langle v, u\rangle}$
  - $\langle u, u\rangle \geq 0$, and $\langle u, u\rangle = 0$ if and only if $u = 0$

[1] kind of a joke, as it means "$1\frac{1}{2}$ linear"

norm and adjoint vector

- the inner product is *conjugate-linear* in its first position:
  - $\langle u+v, w\rangle = \overline{\langle w, u+v\rangle} = \overline{\langle w, u\rangle} + \overline{\langle w, v\rangle} = \langle u, w\rangle + \langle v, w\rangle$
  - $\langle \alpha u, v\rangle = \overline{\langle v, \alpha u\rangle} = \overline{\alpha}\overline{\langle v, u\rangle} = \overline{\alpha}\,\langle u, v\rangle$
- **definition.** an inner product $\langle\cdot,\cdot\rangle$ induces a *norm* $\|\cdot\| : \mathbb{C}^n \to \mathbb{R}$:

$$\|u\| = \sqrt{\langle u, u\rangle}$$

- the *hermitian transpose (adjoint)* of a vector $v \in \mathbb{C}^n$ is the row vector

$$v^* = [\overline{v_1}, \dots, \overline{v_n}]$$

- the usual inner product is just a matrix product on $\mathbb{C}^n$:

$$\langle u, v\rangle = u^* v = \begin{bmatrix}\overline{u_1} & \cdots & \overline{u_n}\end{bmatrix}\begin{bmatrix}v_1 \\ \vdots \\ v_n\end{bmatrix}$$

linear-dependence, span, basis

- **definition.** a (finite) set of vectors $\{v_i\}_{i=1}^m \subset \mathbb{C}^n$ is *linearly-dependent* if there exist scalars $\alpha_i \in \mathbb{C}$, not all zero, so that $\alpha_1 v_1 + \dots + \alpha_m v_m = 0$
  - a set of vectors is *linearly-independent* if it is not linearly-dependent
  - vectors in a linearly-independent set are nonzero
- **definition.** a finite set of vectors $\{v_i\}_{i=1}^m$ *span* $\mathbb{C}^n$ if for any $w \in \mathbb{C}^n$ there exist scalars α_i so that $w = \alpha_1 v_1 + \dots + \alpha_m v_m$
- **lemma.**[2]
  - if $\{v_i\}_{i=1}^m \subset \mathbb{C}^n$ is linearly-independent then $m \le n$
  - if $\{v_i\}_{i=1}^m \subset \mathbb{C}^n$ spans $\mathbb{C}^n$ then $m \ge n$
- **definition.** a finite set of vectors $\{v_i\}_{i=1}^m \subset \mathbb{C}^n$ is a *basis* if the set is linearly-independent and it spans $\mathbb{C}^n$
  - by the lemma, $m = n$
  - the *dimension* n is well-defined as the number of elements in a basis

[2] https://en.wikipedia.org/wiki/Steinitz_exchange_lemma

linear operators and matrices

- **definition.** $A : \mathbb{C}^n \to \mathbb{C}^n$ is *linear* if $A(\alpha u + \beta v) = \alpha A(u) + \beta A(v)$
  - we call such a function a *linear operator* and we write $Au = A(u)$
- given a basis, one may represent a linear operator as a (square) *matrix*
  - matrix multiplication is just function composition
- **definition.** A is *invertible* if there exists B so that $AB = BA = I$
- **lemma.** matrix A is invertible if and only if its columns form a basis

meanings of matrix multiplication

- the "purpose" of a matrix $A \in \mathbb{C}^{n\times m}$ is to multiply vectors $u \in \mathbb{C}^m$
  - a matrix is merely the representation of a linear operator
  - operations on entries, e.g. $\operatorname{det}()$ or row operations, are less fundamental
- when you see "Au" remember two views:

$$Au = \begin{bmatrix} r_1 \\ \hline r_2 \\ \hline \vdots \\ \hline r_n \end{bmatrix} u = \begin{bmatrix} r_1 u \\ r_2 u \\ \vdots \\ r_n u \end{bmatrix}$$

$$Au = \left[\begin{array}{c|c|c|c} a_1 & a_2 & \dots & a_m \end{array}\right] \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} = u_1 \begin{bmatrix} a_1 \end{bmatrix} + u_2 \begin{bmatrix} a_2 \end{bmatrix} + \dots + u_m \begin{bmatrix} a_m \end{bmatrix}$$

  - A acts on u acts from the left; each entry is an inner product
  - u acts from the right on A; get a linear-combination of columns
- also: if $A \in \mathbb{C}^{n\times n}$ is invertible then $A^{-1}w$ computes the coefficients of w in the basis of columns of A

- **definition.** the *hermitian transpose* or *adjoint* of $A \in \mathbb{C}^{n\times m}$ is $A^* \in \mathbb{C}^{m\times n}$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & & a_{2m} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \quad \to \quad A^* = \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \dots & \overline{a_{n1}} \\ \overline{a_{12}} & \overline{a_{22}} & & \overline{a_{n2}} \\ \vdots & & \ddots & \vdots \\ \overline{a_{1m}} & \overline{a_{2m}} & \dots & \overline{a_{nm}} \end{bmatrix}$$

- **definition.** a basis $\{v_i\}_{i=1}^n$ of $\mathbb{C}^n$ is *orthonormal* (ON) if

$$\langle v_i, v_j \rangle = v_i^* v_j = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

  - "ortho" means $\langle v_i, v_j \rangle = 0$ if $i \neq j$ and "normal" means $\|v_i\| = 1$
- **definition.** a matrix U is *unitary* if $U^* U = I$, thus $U^{-1} = U^*$
- **lemma.** a matrix U is unitary if and only if its columns form an ON basis

proof. The entries of a matrix product are inner products between the rows of the left factor and the columns of the right factor. The entries of I are δ_{ij}. $\square$

Gram-Schmidt process

- given a set of vectors $\{w_i\}_{i=1}^m \subset \mathbb{C}^n$ we can generate new orthonormal vectors which span the same subspace
  - if the original set spans $\mathbb{C}^n$ then the result is an ON basis
- formulas:

$$
\begin{aligned}
\tilde v &= w_1 \quad \to & v_1 &= \tilde v / \|\tilde v\| \\
\tilde v &= w_2 - \langle v_1, w_2 \rangle\, v_1 \quad \to & v_2 &= \tilde v / \|\tilde v\| \\
\tilde v &= w_3 - \langle v_1, w_3 \rangle\, v_1 - \langle v_2, w_3 \rangle\, v_2 \quad \to & v_3 &= \tilde v / \|\tilde v\| \\
\tilde v &= w_4 - \langle v_1, w_4 \rangle\, v_1 - \langle v_2, w_4 \rangle\, v_2 - \langle v_3, w_4 \rangle\, v_3 \quad \to & v_4 &= \tilde v / \|\tilde v\| \\
&\vdots & &\vdots
\end{aligned}
$$

  - exception: if $\tilde v = 0$ at any stage then we ignore that w_i and skip to w_{i+1}
  - notice the triangular structure
- **lemma.**
  - $\{v_i\}$ is ON
  - $\mathrm{span}\{v_i\} = \mathrm{span}\{w_i\}$
  - if $\{w_i\}_{i=1}^m$ spans $\mathbb{C}^n$, thus $m \geq n$, then $\{v_i\}_{i=1}^n$ is an ON basis

Gram-Schmidt is QR

- the Gram-Schmidt formulas are of the form
$$\alpha_i v_i \overset{*}{=} w_i - \sum_{j=1}^{i-1} \beta_{ji} v_j$$
where $\alpha_i = \|\tilde{v}\|$ is a normalization constant and $\beta_{ji} = \langle v_j, w_i \rangle$
- moving the v_i to the left in $*$, and writing vectors as columns gives
$$\begin{bmatrix} v_1 \,\Big|\, v_2 \,\Big|\, \dots \,\Big|\, v_m \end{bmatrix} \begin{bmatrix} \alpha_1 & \beta_{12} & \beta_{13} & \dots & \beta_{1m} \\ 0 & \alpha_2 & \beta_{23} & & \beta_{2m} \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \dots & & \alpha_m \end{bmatrix} = \begin{bmatrix} w_1 \,\Big|\, w_2 \,\Big|\, \dots \,\Big|\, w_m \end{bmatrix}$$
- this is a “reduced” QR decomposition
$$\hat{Q}\hat{R} = A$$
  - A is $n \times m$ and contains the original vectors w_i,
  - $\hat{Q}$ is the same size as A and contains the ON vectors v_i,
  - and $\hat{R}$ is (upper-) right-triangular and $m \times m$, thus small if $m \ll n$
- if $m = n$ and columns of A span $\mathbb{C}^n$ (A has *full rank*) then Q is unitary
- methods: numerical QR uses Householder, not Gram-Schmidt, for stability reasons

Gram-Schmidt process: example 1

- suppose we have $m = 3$ vectors in $\mathbb{C}^3$:

$$w_1 = \begin{bmatrix} 9 \\ 3 \\ 4 \end{bmatrix}, \quad w_2 = \begin{bmatrix} 1 \\ 6 \\ 9 \end{bmatrix}, \quad w_3 = \begin{bmatrix} 6 \\ 7 \\ 3 \end{bmatrix}$$

- applying the formulas on slide 10:

$$v_1 = \begin{bmatrix} 0.87416 \\ 0.29139 \\ 0.38851 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -0.48456 \\ 0.46984 \\ 0.73787 \end{bmatrix}, \quad v_3 = \begin{bmatrix} -0.03247 \\ 0.83327 \\ -0.55191 \end{bmatrix}$$

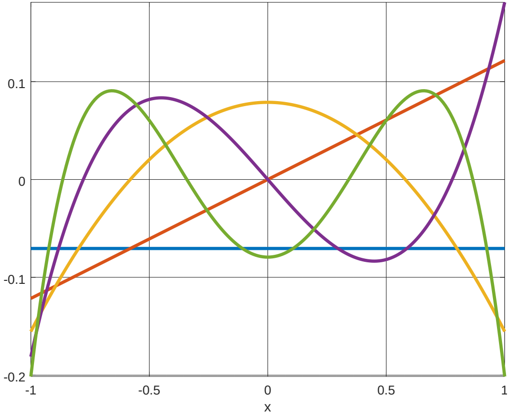
- compare this MATLAB calculation:

```
>> A = [9 1 6; 3 6 7; 4 9 3];
>> [Q,R] = qr(A)
Q =
    -0.87416     0.48456    0.032465
    -0.29139    -0.46984    -0.83327
    -0.38851    -0.73787     0.55191
R =
     -10.296     -6.1191     -8.4502
           0     -8.9753     -2.5952
           0           0     -3.9824
```

Gram-Schmidt process: example 2

- what is this MATLAB calculation doing?:

```
>> x = (-1:.01:1)';
>> A = [x.^0 x.^1 x.^2 x.^3 x.^4];
>> size(A)
ans =
      201        5
>> [Q,R] = qr(A,0);
>> size(Q)
ans =
      201        5
>> plot(x,Q),  xlabel x
>> axis tight,  grid on
```



- the plot shows Legendre polynomials up to degree 4

technique: extend to an ON basis

- a key technique, for proofs related to spectral theory, is to extend $m < n$ ON vectors to an ON basis
- let $\{e_i\}_{i=1}^n$ be the standard basis of $\mathbb{C}^n$, with $(e_i)_j = \delta_{ij}$
- **method.** given ON vectors $\{u_1, \dots, u_m\}$, for $1 \le m < n$, apply the Gram-Schmidt process to

$$w_1 = u_1, \quad \dots \quad w_m = u_m, \quad w_{m+1} = e_1, \quad \dots \quad w_{m+n} = e_n$$

  - note: this set of $m + n$ vectors does indeed span $\mathbb{C}^n$!
  - the first m steps of G.-S. are trivial, but after that there will be discarded vectors because $\|\tilde{v}\| = 0$
  - result is ON set $\{v_i\}_{i=1}^n$ where first m vectors were given
- as matrices, if $\hat{Q}$ has ON columns then we extend to a unitary matrix:

$$\hat{Q} = \left[\begin{array}{c|c|c} u_1 & \dots & u_m \end{array}\right] \quad \to \quad Q = \left[\begin{array}{c|c|c|c|c|c} u_1 & \dots & u_m & v_{m+1} & \dots & v_n \end{array}\right]$$

eigenvalues and eigenvectors

- **definition.** given a square matrix $A \in \mathbb{C}^{n\times n}$, $v \in \mathbb{C}^n$ is an *eigenvector* if $v \neq 0$ and if there exists $\lambda \in \mathbb{C}$ so that

$$Av = \lambda v$$

  - λ is the *eigenvalue* for v
  - idea: A acts in a simple way on (multiples of) v, simply by scaling
- **definition.** the *spectrum* $\sigma(A)$ of $A \in \mathbb{C}^{n\times n}$ is the set of its eigenvalues
- **lemma.** $\lambda \in \sigma(A)$ if and only if $\det(\lambda I - A) = 0$

proof. $\lambda \in \sigma(A) \iff \exists$ nonzero soln. to $(\lambda I - A)v = 0 \iff \det(\lambda I - A) = 0$ $\square$

- **corollary.** $\sigma(A)$ is nonempty

proof. $p(\lambda) = \det(\lambda I - A)$ is a degree n polynomial, which has a root $\lambda \in \mathbb{C}$ $\square$

hermitian matrices and their eigenvalues

- general facts about adjoints:
  - $(AB)^* = B^* A^*$
  - in usual inner product $\langle v, w\rangle = v^* w$: $\langle v, Aw\rangle = \langle A^* v, w\rangle$
- **definition.** $A \in \mathbb{C}^{n\times n}$ is *hermitian* if $A^* = A$
  - also called *self-adjoint*
  - $\overline{a_{ij}} = a_{ji}$ . . . so the diagonal entries of A are real
  - in usual inner product $\langle v, w\rangle = v^* w$: $\langle v, Aw\rangle = \langle Av, w\rangle$
  - if A has real entries then $A^\top = A$, and A is *symmetric*
- **lemma.** if A is hermitian and $\lambda \in \sigma(A)$ then λ is real

proof. a classic exercise . . . for you

- **lemma.** if A is hermitian and $v, w \in \mathbb{C}^n$ are eigenvectors associated to distinct eigenvalues then v, w are orthogonal

proof. a classic exercise: if $Av = \lambda v$ and $Aw = \mu w$ then λ, μ are real so

$$(\lambda-\mu)\langle v, w\rangle = \langle \lambda v, w\rangle - \langle v, \mu w\rangle = \langle Av, w\rangle - \langle v, Aw\rangle = \langle v, Aw\rangle - \langle v, Aw\rangle = 0,$$

and we have assumed $\lambda - \mu \neq 0$ □

eigendecompositions

- for any $A \in \mathbb{C}^{n\times n}$, if $v_1, \dots, v_m$ are eigenvectors with eigenvalues $\lambda_1, \dots, \lambda_m$ then we can collect the statements $Av_i = \lambda_i v_i$ as

$$A \left[\begin{array}{c|c|c} v_1 & \dots & v_m \end{array}\right] = \left[\begin{array}{c|c|c} v_1 & \dots & v_m \end{array}\right] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{bmatrix}$$

- equivalently:

$$AV = V\Lambda$$

- if V is square and invertible then we have a decomposition of A:

$$A = V\Lambda V^{-1}$$

- why does this matter? often A is iterated, or a polynomial is applied,

$$A^k = (V\Lambda V^{-1})^k = V\Lambda V^{-1} V\Lambda V^{-1} \dots V\Lambda V^{-1} = V\Lambda^k V^{-1}$$

$$p(A) = V\,p(\Lambda)\,V^{-1}$$

  - Λ^k, $p(\Lambda)$ are easy-to-understand diagonal matrices

similarity and diagonalizability

- **definition.** A, B are *similar* if there exists V invertible so that
$$A = VBV^{-1}$$
- **lemma.** if A, B are similar then $\sigma(A) = \sigma(B)$

proof. noting $\det(TS) = \det(T)\det(S)$, we have
$$\begin{aligned}\det(\lambda I - A) &= \det(\lambda V V^{-1} - VBV^{-1}) = \det(V)\det(\lambda I - B)\det(V^{-1}) \\ &= \det(V)\det(\lambda I - B)\det(V)^{-1} = \det(\lambda I - B) \quad \square\end{aligned}$$

- **lemma.** if A is triangular or diagonal then $\sigma(A) = \{a_{ii}\}$ (diagonal entries)

proof. $\det(\lambda I - A) = \prod_{i=1}^{n} \lambda - a_{ii}$ $\square$

- **definition.** A is *diagonalizable* if there exists V invertible and Λ diagonal so that
$$A = V\Lambda V^{-1}$$
  - equivalently, A is diagonalizable if it is similar to a diagonal matrix
  - not every square matrix is diagonalizable!

non-diagonalizable ("defective") matrices

- if you pick a matrix $A \in \mathbb{C}^{n\times n}$ at random then it will be diagonalizable with probability one,[3] but non-diagonalizable matrices exist
- simplest example (*check all this!*): $\quad A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$
    - note A is triangular, and $\sigma(A) = \{0\}$
    - if $A = V\Lambda V^{-1}$ for V invertible then columns of V would be linearly-independent eigenvectors: $Av_1 = 0v_1$ and $Av_2 = 0v_2$
    - but in fact $(\lambda I - A)v = 0$ has only a one-dimensional solution space
- all non-diagonal examples A are built by choosing J to be a "non-diagonal Jordan form," with at least one block of this form on the diagonal:

$$\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}, \quad \dots,$$

and with J otherwise diagonal, and defining $A = VJV^{-1}$ for some invertible V

[3] this assumes only that your probability measure has a continuous density with respect to Lebesgue measure

eigenvalue-revealing decompositions

- there are four famous "eigenvalue-revealing" decompositions of square matrices $A \in \mathbb{C}^{n\times n}$:

$A = V\Lambda V^{-1}$	if A is diagonalizable
$A = VJV^{-1}$	for any A, where J is special triangular *Jordan form*
$A = Q\Lambda Q^*$	if A is hermitian, where Q is unitary
$A = QTQ^*$	for any A, where T is triangular and Q is unitary

- respectively:

	MATLAB commands
$A = V\Lambda V^{-1}$ is *eigendecomposition* or *diagonalization*	`eig(A)`
$A = VJV^{-1}$ is *Jordan canonical form*	
$A = Q\Lambda Q^*$ is *spectral theorem*	`eig(A)`
$A = QTQ^*$ is *Schur decomposition*	`schur(A,'complex')`

- the Jordan canonical form cannot be computed if rounding errors exist[4]
- the Schur decomposition is the most important in practice, as it always exists and it can be stably computed over $\mathbb{R}$ or $\mathbb{C}$

[4] G. Golub & J. Wilkinson 1976, *Ill-conditioned eigensystems and the computation of the Jordan canonical form*, SIAM Review 18(4), 578–619

Schur decomposition

- **theorem.** if $A \in \mathbb{C}^{n\times n}$ then there exist $T, Q \in \mathbb{C}^{n\times n}$, with T upper-triangular and Q unitary, so that

$$A = QTQ^*$$

proof. Induct on n. If $n = 1$ then the result follows with $T = A = [a_{11}]$ and $Q = I = [1]$.
For $n > 1$ let $v \neq 0$ be an eigenvector of A, with eigenvalue λ. Let $u_1 = v/\|v\|$ and extend to a unitary matrix:

$$U = \left[\begin{array}{c|c|c|c} u_1 & u_2 & \dots & u_n \end{array}\right]$$

Apply A and note that $Au_1 = \lambda u_1$:

$$AU = \left[\begin{array}{c|c|c|c} \lambda_1 u_1 & Au_2 & \dots & Au_n \end{array}\right]$$

Apply $U^* = U^{-1}$, observe that $U^*Au_1 = \lambda_1 U^* u_1 = \lambda_1 e_1$, and write $w_j = U^*Au_j$:

$$U^*AU = \left[\begin{array}{c|c|c|c} \lambda_1 e_1 & w_2 & \dots & w_n \end{array}\right]$$

(We don't care about the form of the vectors $w_2, \dots, w_n \in \mathbb{C}^n$.)

Schur decomposition, proof cont.

We have made progress toward upper-triangular form. In fact, we may write

$$U^* A U = \left[\begin{array}{c|c} \lambda_1 & z^* \\ \hline 0 & M \end{array}\right]$$

for some $z \in \mathbb{C}^{n-1}$ and $M \in \mathbb{C}^{n-1 \times n-1}$. By induction, M has a Schur decomposition, $M = \hat{Q}\hat{T}\hat{Q}^*$, where $\hat{T}$, $\hat{Q}$ are the same size as M. Note that

$$\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \hat{Q} \end{array}\right]$$

is unitary. Now we can transform the whole matrix $U^* A U$ to triangular form:

$$\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \hat{Q}^* \end{array}\right] U^* A U \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \hat{Q} \end{array}\right] = \left[\begin{array}{c|c} \lambda_1 & z^* \hat{Q} \\ \hline 0 & \hat{T} \end{array}\right]$$

Now let

$$T = \left[\begin{array}{c|c} \lambda_1 & z^* \hat{Q} \\ \hline 0 & \hat{T} \end{array}\right], \quad Q = U \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \hat{Q} \end{array}\right]$$

We have $A = Q^* T Q$. $\square$

- note key steps at start of proof: "let $v \neq 0$ be an eigenvector of A" and "extend to a unitary matrix"

Schur decomposition: computed examples

- note the diagonal entries of T contain the eigenvalues (compare `eig(A)`)
- example 1: general case

```
>> A=randn(4,4);
>> [Q,T] = schur(A)
Q =
   0.25290   0.78457  -0.32111   0.46624
   0.12587   0.07899  -0.70892  -0.68945
  -0.82852   0.47758   0.15009  -0.25088
  -0.48348  -0.38746  -0.60974   0.49431
T =
   1.94208  -1.22908  -1.37600  -0.70166
   0.00000  -1.81581   0.21700  -0.51769
   0.00000   0.00000   0.69477  -1.15766
   0.00000   0.00000   0.00000  -0.59708
>> norm(A-Q*T*Q')
ans =    2.0928e-15
```

  - we just got lucky; see `help schur` for real vs. complex Schur decompositions

- example 2: hermitian case

```
>> B = randn(4,4);  A = B + B';
>> [Q,T] = schur(A);  T
T =
   3.42692  -0.00000  -0.00000  -0.00000
   0.00000   0.96497   0.00000  -0.00000
   0.00000   0.00000  -2.92548   0.00000
   0.00000   0.00000   0.00000  -4.61828
```

normal matrices

- **lemma.** if A is hermitian then the Schur decomposition is a unitary diagonalization

proof. $A^* = A$ so $QT^*Q^* = QTQ^*$ so $T^* = T$ $\square$

- there is a larger class of matrices where this happens
- **definition.** $A \in \mathbb{C}^{n\times n}$ is *normal* if $AA^* = A^*A$
- examples:
  - A hermitian $\implies$ A normal
  - U unitary $\implies$ U normal
  - S skew-hermitian[5] $\implies$ S normal

[5] $S^* = -S$

- **corollary (spectral theorem).** if A is normal then there exists Λ diagonal and Q unitary so that
$$A = Q\Lambda Q^*$$

proof. From the Schur decomposition, $A = QTQ^*$. Since A is normal, it follows that $TT^* = T^*T$. But T is upper-triangular, so

$$T = \left[\begin{array}{c|c} t_{11} & z^* \\ \hline 0 & R \end{array}\right]$$

where R is also upper triangular. An easy calculation shows $(TT^*)_{11} = |t_{11}|^2 + \|z\|^2$ while $(T^*T)_{11} = |t_{11}|^2$. Thus $z = 0$. Now induct to show T is diagonal. $\square$

where we stand

- in part II we will discuss consequences of the spectral theorem
- …and get to the singular value decomposition
- almost everything we do will have some kind of analog in ∞-dimensions
- …and appear somewhere in our textbook[6]
- …but most proof steps do not extend directly to ∞-dimensions
- **questions.** in an ∞-dimensional Hilbert space,
  - what is the meaning of “span” and “basis”?
  - are matrices meaningful?
  - is a one-to-one linear operator invertible?
  - does the Gram-Schmidt process work as before?
  - does every linear operator have an eigenvector?
  - is there a Schur decomposition of every linear operator?
  - is there a spectral theorem of hermitian or normal operators?

[6] Muscat, *Functional Analysis*; see Chapter 10 for Hilbert spaces