Sergio Garcia Tapia
Numerical Linear Algebra, Lloyd Trefethen and David Bau III
Lecture 2: Orthogonal Vectors and Matrices
June 19, 2024

Lecture 2: Orthogonal Vectors and Matrices

Exercise 1. Show that if a matrix A is both triangular and unitary, then it is diagonal.

Solution:

Proof 1:

Proof. Suppose A is an $m \times m$ unitary and upper-triangular matrix. Since A is unitary, its columns are orthonormal. If e_j is the j-th standard basis vectors of $\mathbb{C}^m$, then $a_j = Ae_j$, where a_j is the j-th column of A. Since A is upper triangular, we see that the j-th column is 0 beyond the j-th entry. Hence,

$$a_j = Ae_j = \sum_{k=1}^{j} c_k e_k$$

Since A is unitary, we know that $a_i^* a_j = \delta_{ij}$, where δ_{ij} is the Kronecker delta, with value 1 when $i = j$ and 0 when $i \neq j$. Thus

$$\delta_{ij} = a_i^* a_j = \sum_{k=1}^{j} c_k a_i^* a_k$$

by the bilinearity of the inner product. If we take $i < j$, we get $0 = c_i$. Thus, $Ae_j = c_j e_j$, and hence A is a diagonal matrix. □

Proof 2:

Proof. Suppose A is an $m \times m$ upper-triangular and unitary matrix. Since A is unitary, it follows that the adjoint of A is its inverse, meaning $A^* = A^{-1}$. Thus if a_i is the i-th column of A and a_i^* is its adjoint, their inner product is $a_i^* a_j = \delta_{ij}$. The δ_{ij} stands for the Kronecker delta, whose value is 1 if $i = j$ and 0 otherwise.

We already know that $a_{ij} = 0$ for $i > j$, so we have to show that $a_{ij} = 0$ for $i < j$. The proof is by induction on the column index of j. Consider the first column ($j = 1$). Since A is upper-triangular, it follows that $a_{k1} = 0$ for $k > 1$, and there no entries with $k < 1$. Moreover, the fact that A is unitary means that its columns are orthonormal, so

$$1 = a_1^* a_1 = \overline{a_{11}} a_{11}$$

and hence $a_{11} \neq 0$. If $j = 2$, and note that because the columns of A are orthogonal, we have

$$0 = a_1^* a_2 = \sum_{k=1}^{m} \overline{a_{k1}} a_{k2} = \overline{a_{11}} a_{12}$$

where the sum collapsed because $a_{k1} = 0$ for $k > 1$. Since $\overline{a_{11}} \neq 0$, we conclude that $a_{12} = 0$, and hence all entries in a_2 except a_{22} are 0.

Suppose that $j > 1$ and that $1 \leq i < j$. Then by our induction hypothesis, the k-th entry of the i-th column a_i is 0 if $k \neq i$. Since $i \neq j$, a_i and a_j are orthogonal, so

$$0 = a_i^* a_j = \sum_{k=1}^{m} \overline{a_{ki}} a_{kj} = \overline{a_{ii}} a_{ij}$$

Since $a_i^* a_i = 1$, we know that $a_{ii} \neq 0$, so we conclude that $a_{ij} = 0$. Since we also know that $a_{ij} = 0$ for $i > j$ due to the fact that A is upper-triangular, we conclude that $a_{ij} = 0$ for $i \neq j$. This holds by induction on j, and hence A is diagonal. □

Exercise 2. The Pythagorean Theorem asserts that for a set of n orthogonal vectors $\{x_i\}$,

$$\left\| \sum_{i=1}^{n} x_i \right\|^2 = \sum_{i=1}^{n} \|x_i\|^2$$

(a) Prove this in the case $n = 2$ by explicit computation of $\|x_1 + x_2\|^2$.

(b) Show that this computation also establishes the general case, by induction.

Solution:

(a) *Proof.* If x_1 and x_2 are orthogonal, then their inner product is $x_1^* x_2 = 0$. Meanwhile, the notation $\|x_i\|^2$ refers to the squared norm, or the value $x_i^* x_i$. Thus, by the bilinearity of the inner product, we have

$$\begin{aligned}
\|x_1 + x_2\|^2 &= (x_1 + x_2)^*(x_1 + x_2) \\
&= (x_1 + x_2)^* x_1 + (x_1 + x_2)^* x_2 \\
&= x_1^* x_1 + x_2^* x_1 + x_1^* x_2 + x_2^* x_2 \\
&= \|x_1\|^2 + 0 + 0 + \|x_2\|^2 \\
&= \|x_1\|^2 + \|x_2\|^2
\end{aligned}$$

□

(b) *Proof.* The proof is by induction on n, the size of the orthogonal set $\{x_i\}$. The case with 1 vector holds trivially, and the case with 2 vectors has been shown in (a). Suppose that $n > 1$ and that all orthogonal sets with less than n vectors satisfy the given equality. Then

$$
\begin{aligned}
\left\|\sum_{i=1}^{n} x_i\right\|^2 &= \left\|x_n + \sum_{i=1}^{n-1} x_i\right\|^2 \\
&= \left(x_n + \sum_{i=1}^{n-1} x_i\right)^* \left(x_n + \sum_{i=1}^{n-1} x_i\right) \\
&= x_n^* x_n + x_n^* \left(\sum_{i=1}^{n-1} x_i\right) + \left(\sum_{i=1}^{n-1} x_i\right)^* x_n + \left(\sum_{i=1}^{n-1} x_i\right)^* \left(\sum_{i=1}^{n-1} x_i\right) \\
&= \|x_n\|^2 + \sum_{i=1}^{n-1} x_n^* x_i + \sum_{i=1}^{n-1} x_i^* x_n + \left\|\sum_{i=1}^{n-1} x_i\right\|^2 \\
&= \|x_n\|^2 + \sum_{i=1}^{n-1}(0) + \sum_{i=1}^{n-1}(0) + \sum_{i=1}^{n-1}\|x_i\|^2 \\
&= \|x_n\|^2 + \sum_{i=1}^{n-1}\|x_i\|^2 \\
&= \sum_{i=1}^{n}\|x_i\|^2
\end{aligned}
$$

□

Exercise 3. Let $A \in \mathbb{C}^{m\times m}$ be hermitian. An eigenvector of A is a nonzero vector $x \in \mathbb{C}^{m\times m}$ such that $Ax = \lambda x$ for some $\lambda \in \mathbb{C}$, the corresponding eigenvalue.

(a) Prove that all eigenvalues of A are real.

(b) Prove that if x and y are eigenvectors corresponding to distinct eigenvalues, then x and y are orthogonal.

Solution:

(a) *Proof.* Suppose λ is an eigenvalue of A and let v be its eigenvector. Since A is hermitian, we know that $A^* = A$. Since $(Av)^* = v^*A^*$, we have

$$
\begin{aligned}
(Av)^*v &= v^*A^*v = v^*Av = v^*(\lambda v) = \lambda\|v\|^2 \\
(Av)^*v &= (\lambda v)^*v = \bar{\lambda}v^*v = \bar{\lambda}\|v\|^2
\end{aligned}
$$

Since these two quantities are equal, we are led to $(\lambda - \bar{\lambda})\|v\|^2 = 0$. Since v is an eigenvector, we know that $\|v\| \neq 0$, so we conclude $(\lambda - \bar{\lambda}) = 0$, and hence $\lambda = \bar{\lambda}$, implying λ is real. □

(b) *Proof.* Suppose x and y are eigenvectors of A corresponding to eigenvalues λ and μ, respectively. Then $Ax = \lambda x$ and $Ay = \mu y$. Now

$$
\begin{aligned}
(Ax)^*y &= (\lambda x)^*y = \bar{\lambda}x^*y = \lambda x^*y \\
(Ax)^*y &= x^*A^*y = x^*Ay = x^*\bar{\mu}y = \mu x^*y
\end{aligned}
$$

□

These two quantities are equal, so $(\lambda - \mu)x^*y = 0$. Since $\lambda \neq \mu$, we conclude that $x^*y = 0$, and hence, x and y are orthogonal.

Exercise 4. What can be said about the eigenvalues of a unitary matrix?

Solution: Suppose that A is an $m \times m$ unitary matrix, and v is an eigenvalue of A with eigenvalue λ, so that $Av = \lambda v$. Since A is unitary, it preserves norms, meaning that $\|Ax\| = \|x\|x$ for every $x \in \mathbb{C}^m$, so

$$\|v\| = \|Av\| = \|\lambda v\| = |\lambda| \cdot \|v\|$$

Since $v \neq 0$, we can divide by it and conclude that $|\lambda| = 1$. Thus, every eigenvalue of $\mathbb{A}$ has absolute value 1, and thus it lies on the unit circle in $\mathbb{C}$.

Exercise 5. Let $S \in \mathbb{C}^{m \times m}$ be *skew-hermitian*, i.e., $S^* = -S$.

(a) Show by using Exercise 2.3 that eigenvalues of S are pure imaginary.

(b) Show that $I - S$ is nonsingular.

(c) Show that the matrix $Q = (I - S)^{-1}(I + S)$, known as the *Caley transform of* S, is unitary (This is a matrix analogue of a linear fractional transformation $(1 + s)/(1 - s)$, which maps the left half of the complex s-plane conformally onto the unit disk).

Solution:

(a) *Proof.* If λ is an eigenvalue of S, then there is a nonzero vector $v \in \mathbb{C}^m$ such that $Sv = \lambda v$. Then

$$(Sv)^*v = v^*S^*v = v^*(-Sv) = v^*(-\lambda v) = -\lambda\|v\|^2$$
$$(Sv)^*v = (\lambda v)^*v = \bar{\lambda}v^*v = \bar{\lambda}\|v\|^2$$

Equating the two, we get $\bar{\lambda}\|v\|^2 = -\lambda\|v\|^2$. Since $v \neq 0$, we have $\|v\| \neq 0$, so dividing by it gives $\bar{\lambda} = -\lambda$. Thus, either $\lambda = 0$, or λ is pure imaginary. □

(b) *Proof.* Suppose there is $v \in V$ such that $(I - S)v = 0$. Then $Sv = v$. If $v \neq 0$, then this implies $\lambda = 1$ is an eigenvalue of S. Since S is skew-symmetric, this would imply that $\bar{\lambda} = -\lambda$, which is impossible since $\lambda = 1$ is real. Thus we in fact have $v = 0$, which means null $(I - S) = \{0\}$. By Theorem 1.3, we conclude $I - S$ is nonsingular. □

(c) *Proof.* A similar argument to (b) shows that $I + S$ is nonsingular. Specifically, if $(I + S)v = 0$ with $v \neq 0$, then $Sv = -v$, implying that $\lambda = -1$ is an eigenvalue of S, again contradicting (a) because S is skew-symmetric so we should have $\bar{\lambda} = -\lambda$. The contradiction implies that $v = 0$, so null $(I + S) = \{0\}$, and hence $I + S$ is invertible.

Now using the fact that $(A^*)^{-1} = (A^{-1})^*$, we have

$$\begin{aligned}
Q^* &= [(I-S)^{-1}(I+S)]^* \\
&= (I+S)^*[(I-S)^{-1}]^* \\
&= (I+S)^*[(I-S)^*]^{-1} \\
&= (I^*+S^*)(I^*-S^*)^{-1} \\
&= (I-S)(I+S)^{-1}
\end{aligned}$$

Moreover, although matrix multiplication is not commutative in general, the matrices $I+S$ and $I-S$ do commute:

$$\begin{aligned}
(I-S)(I+S) &= I^2 + I\cdot S - S\cdot I - S^2 = I^2 - S^2 \\
(I+S)(I-S) &= I^2 - I\cdot S + S\cdot I - S^2 = I^2 - S^2
\end{aligned}$$

Thus, using the fact that $(AB)^{-1} = B^{-1}A^{-1}$, we have

$$\begin{aligned}
Q^*Q &= [(I-S)(I+S)^{-1}](I-S)^{-1}(I+S) \\
&= (I-S)[(I+S)^{-1}(I-S)^{-1}](I+S) \\
&= (I-S)[(I-S)(I+S)]^{-1}(I+S) \\
&= (I-S)[(I+S)(I-S)]^{-1}(I+S) \\
&= (I-S)(I-S)^{-1}(I+S)^{-1}(I+S) \\
&= I\cdot I \\
&= I
\end{aligned}$$

Hence $Q^* = Q^{-1}$, implying Q is unitary. □

Exercise 6. If u and v are m-vectors, the matrix $A = I + uv^*$ is known as a *rank-one perturbation of the identity*. Show that if A is nonsingular, then its inverse has the form $A^{-1} = I + \alpha uv^*$ for some scalar α, and give an expression for α. For what u and v is A singular? If it is singular, what is null (A)?

Solution:

Proof. Suppose A were nonsingular and that its inverse was $I + \alpha uv^*$. Since $AA^{-1} = I$, we get

$$\begin{aligned}
I &= (I+uv^*)(I+\alpha uv^*) \\
&= I\cdot I + I\cdot \alpha uv^* + uv^*\cdot I + \alpha uv^*uv^* \\
&= I + \alpha uv^* + uv^* + \alpha u(v^*u)v^* \\
&= I + \alpha uv^* + uv^* + (\alpha v^*u)uv^*
\end{aligned}$$

Subtracting I on both sides and rearranging, we get

$$0 = (\alpha + 1 + \alpha v^*u)uv^*$$

If u and v are nonzero, then $\alpha + 1 + \alpha v^* u = 0$. If $v^* u \neq -1$, then

$$\alpha(1 + v^* u) = -1$$
$$\alpha = -\frac{1}{1 + v^* u}$$

If $u^* v = -1$, then A is singular. Suppose we had $Aw = 0$ for some $w \in \mathbb{C}^m$. Then $0 = (I + uv^*)w$, so $0 = w + uv^* w$. Then

$$w = -(v^* w)u$$

That is, $w \in \text{span}(u)$, so null $(A) = \text{span}(u)$ Indeed:

$$(I + uv^*)u = u + uv^* u = u + u \cdot (-1) = 0$$

□

Exercise 7. A *Hadamard matrix* is a matrix whose entries are all ± 1 and whose transpose is equal to its inverse times a constant factor. It is known that if A is a Hadamard matrix of dimension $m > 2$, then m is a multiple of 4. It is not known, however, whether there is a Hadamard matrix for every m, though examples have been found for all cases $m \leq 424$.

Show that the following recursive description provides a Hadamard matrix of each dimension $m = 2^k$, $k = 0, 1, 2, \ldots$:

$$H_0 = \begin{bmatrix} 1 \end{bmatrix} \qquad H_{k+1} = \begin{bmatrix} H_k & H_K \\ H_k & -H_k \end{bmatrix}$$

Solution:

Proof. To show that each H_k is a Hadamard matrix, we must show that H_k only has 1 or -1 as entries, that it is invertible, and that there is a constant $c \in \mathbb{C}$ such that $H_k^T = c \cdot H_k^{-1}$. From the recursive description, it's fairly easy to see that it only has 1 and -1's as entries.

The proof is by induction on k. If $k = 0$, then $H_0^T = -H_0^{-1}$. If $k = 1$, then

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad H_1^T = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad H_1^{-1} = -\frac{1}{2} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} = \frac{1}{2} H_1^T$$

Suppose now that $k \geq 1$ and H_k is Hadamard. Then H_k is invertible, and $H_k^T = c H_k^{-1}$ for some $c \in \mathbb{C}$. Then

$$H_{k+1} = \begin{bmatrix} H_k & H_k \\ H_k & -H_k \end{bmatrix}$$

The size of H_k is 2^k, so the size of $H_{k+1} = 2^{k+1}$. If i, j is an entry in the top-left H_k matrix, then $i \leq 2^k$ and $j \leq 2^k$. Thus, when swapped, implying that all such entries remain in the top-left after the transpose. Similarly for the bottom-right corner $-H_k$ matrix. If an entry is in the top-right corner matrix, then $i \leq 2^k$, but $j > 2^k$. When swapped due to the transpose, the entry goes to the bottom-left corner, where the row index exceeds 2^k, but the column index does not. Thus,

$$H_{k+1}^T = \begin{bmatrix} H_k^T & H_k^T \\ H_k^T & -H_k^T \end{bmatrix}$$

Let $h_j^{(k)}$ be the j-th column of H_k, and $h_j^{(k+1)}$ be the j-th column of H_{k+1}. Then $h_j^{(k+1)}$ has two copies of $h_j^{(k)}$ stacked, so when we perform the matrix product by computing the dot product, the result for i and j no greater than 2^k is is

$$[h_i^{(k+1)}]^* h_j^{(k+1)} = (h_i^{(k)})^* h_j^{(k)} + (h_i^{(k)})^* h_j^{(k)} = 2c \cdot \delta_{ij}$$

If we now allow $i > 2^k$ and $j \leq 2^{k+1}$, then then $h_i^{(k+1)}$ consists of $h_{i \bmod 2^k}^{(k)}$ followed by $-h_{i \bmod 2^k}^{(k)}$. Thus

$$[h_i^{(k+1)}]^* h_j^{(k+1)} = (h_i^{(k)})^* h_j^{(k)} - (h_i^{(k)})^* h_j^{(k)} = 0$$

Similar arguments lead to

$$H_{k+1}^T H_{k+1} = \begin{bmatrix} 2cI & 0 \\ 0 & 2cI \end{bmatrix}$$

where I is the $2^k \times 2^k$ identity matrix. Hence, $H_{k+1}^T = 2cH^{-1}$. $\square$