Complex Inner Product Spaces

The $\mathbb{C}^n$ spaces

The prototypical (and most important) real vector spaces are the Euclidean spaces $\mathbb{R}^n$. Any study of complex vector spaces will similar begin with $\mathbb{C}^n$. As a set, $\mathbb{C}^n$ contains vectors of length n whose entries are complex numbers. Thus,

$$\begin{bmatrix} 2+i \\ 3-5i \\ i \end{bmatrix} \in \mathbb{C}^3,$$

$(5,-1)$ is an element found *both* in $\mathbb{R}^2$ and $\mathbb{C}^2$ (and, indeed, all of $\mathbb{R}^n$ is found in $\mathbb{C}^n$), and $(0,0,0,0)$ serves as the *zero* element in $\mathbb{C}^4$. Addition and scalar multiplication in $\mathbb{C}^n$ is done in the analogous way to how they are performed in $\mathbb{R}^n$, except that now the scalars are allowed to be nonreal numbers. Thus, to rescale the vector $(3+i, -2-3i)$ by $1-3i$, we have

$$(1-3i)\begin{bmatrix} 3+i \\ -2-3i \end{bmatrix} = \begin{bmatrix} (1-3i)(3+i) \\ (1-3i)(-2-3i) \end{bmatrix} = \begin{bmatrix} 6-8i \\ -11+3i \end{bmatrix}.$$

Given the notation $\overline{3+2i}$ for the complex conjugate $3-2i$ of $3+2i$, we adopt a similar notation when we want to take the complex conjugate simultaneously of all entries in a vector. Thus,

$$\text{if} \quad \mathbf{z} = \begin{bmatrix} 3-4i \\ 2i \\ -2+5i \\ -1 \end{bmatrix}, \qquad \text{then} \qquad \overline{\mathbf{z}} = \begin{bmatrix} 3+4i \\ -2i \\ -2-5i \\ -1 \end{bmatrix}.$$

Both $\mathbf{z}$ and $\overline{\mathbf{z}}$ are vectors in $\mathbb{C}^4$. In general, if the entries of $\mathbf{z}$ are all real numbers, then $\overline{\mathbf{z}} = \mathbf{z}$.

The inner product in $\mathbb{C}^n$

In $\mathbb{R}^n$, the length of a vector $\|\mathbf{x}\| = \sqrt{\mathbf{x}\cdot\mathbf{x}}$ is a real, nonnegative number. The modulus, or length, of a complex number $z = a + ib$ is real and nonnegative as well:

$$|z| = \sqrt{\overline{z}z} = \sqrt{(a-ib)(a+ib)} = \sqrt{a^2+b^2}, \qquad \text{or} \qquad |z|^2 = \overline{z}z.$$

A natural idea, therefore, is to define an inner product between vectors $\mathbf{z} = (z_1, \ldots, z_n)$, $\mathbf{w} = (w_1, \ldots, w_n) \in \mathbb{C}^n$ in this manner:

$$\langle \mathbf{w}, \mathbf{z} \rangle = \sum_{j=1}^{n} \overline{w_j} z_j = \overline{w_1} z_1 + \cdots + \overline{w_n} z_n = \mathbf{w}^{\mathrm{H}}\mathbf{z} = \overline{\mathbf{w}}^{\mathrm{T}}\mathbf{z}. \tag{1}$$

Here, $\mathbf{w}^{\mathrm{H}}$ stands for $\overline{\mathbf{w}}^{\mathrm{T}}$, the **conjugate transpose** of $\mathbf{w}$. For instance,

$$\text{if} \quad \mathbf{z} = \begin{bmatrix} 3-4i \\ 2i \\ -2+5i \\ -1 \end{bmatrix}_{4\times 1} \qquad \text{then} \qquad \mathbf{z}^{\mathrm{H}} = \overline{\mathbf{z}}^{\mathrm{T}} = \begin{bmatrix} 3+4i & -2i & -2-5i & -1 \end{bmatrix}_{1\times 4}.$$

Remarks

- Some authors define the inner product of complex vectors $\mathbf{u}, \mathbf{v}$ to be the conjugate transpose of the second vector multiplied by the first—i.e., $\mathbf{v}^{\mathrm{H}}\mathbf{u}$. The two definitions suit the same purpose, but do not yield the same result. For example, if $\mathbf{u} = (2+i, 1-3i, 8)$ and $\mathbf{v} = (-i, 3+2i, 1-i)$, then
$$\mathbf{v}^{\mathrm{H}}\mathbf{u} = \begin{bmatrix} i & 3-2i & 1+i \end{bmatrix} \begin{bmatrix} 2+i \\ 1-3i \\ 8 \end{bmatrix} = 4-i,$$
but
$$\mathbf{u}^{\mathrm{H}}\mathbf{v} = \begin{bmatrix} 2-i & 1+3i & 8 \end{bmatrix} \begin{bmatrix} -i \\ 3+2i \\ 1-i \end{bmatrix} = 4+i.$$
That is, the result of the one is always the complex conjugate of the other. It is a matter of preference which definition one uses, and I opt for this one just to be consistent with Strang's usage.

- When the entries of $\mathbf{z}, \mathbf{w}$ are all real numbers (that is, $\mathbf{z}, \mathbf{w} \in \mathbb{R}^n$), this new understanding for inner product exactly matches the dot product—that is, $\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{z} \cdot \mathbf{w}$.

- The inner product of vectors in $\mathbb{C}^n$ no longer exclusively produces real numbers, as seen in the example above. However, when taking an inner product of $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ with itself, the result
$$\langle \mathbf{z}, \mathbf{z} \rangle = \mathbf{z}^{\mathrm{H}}\mathbf{z} = \sum_{j=1}^{n} \overline{z_j} z_j = \sum_{j=1}^{n} |z_j|^2,$$
is the sum of the moduli of the components of $\mathbf{z}$, guaranteed to be nonnegative. Thus, we define *length* for vectors $\mathbf{z}$ in $\mathbb{C}^n$ to be
$$\|\mathbf{z}\| = \sqrt{\langle \mathbf{z}, \mathbf{z} \rangle},$$
and note that the only instance in which $\|\mathbf{z}\| = 0$ is when $\mathbf{z}$ is, itself, the zero vector.

The essential list of properties that the inner product in $\mathbb{C}^n$ has, for all vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{C}^n$ and all scalars $\mathbf{c}$, is

(i) $\langle \mathbf{v}, \mathbf{v} \rangle \geqslant 0$.

(ii) $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ implies $\mathbf{v} = \mathbf{0}$.

(iii) $\langle \mathbf{u} + \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle$.

(iv) $\langle c\mathbf{u}, \mathbf{v} \rangle = \bar{c} \langle \mathbf{u}, \mathbf{v} \rangle$.

(v) $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$.

Note that (iii) and (v) together imply that

$$\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle,$$

while (iv) and (v) together give that

$$\langle \mathbf{u}, c\mathbf{v} \rangle = c \langle \mathbf{u}, \mathbf{v} \rangle.$$

Conjugate transpose of a matrix

Suppose, now, that $\mathbf{A}$ is an m-by-n matrix whose entries are complex numbers. The idea of a *conjugate transpose* $\mathbf{A}^{\mathrm{H}}$ makes sense, as it did for vectors in $\mathbb{C}^n$. In fact, it is precisely what is computed when the *prime* symbol is invoked in Octave.

```
octave:127> A = randi(7,3,2)−4*ones(3,2) + (randi (5,3,2) −4*ones(3,2))*i
A =

  −2 − 3i   2 + 1i
   0 − 1i  −1 − 1i
   2 − 2i   1 + 1i

octave:128> A'
ans =

  −2 + 3i   0 + 1i   2 + 2i
   2 − 1i  −1 + 1i   1 − 1i
```

Of course, taking the conjugate transpose of a matrix twice returns one to the original: $(\mathbf{A}^{\mathrm{H}})^{\mathrm{H}} = \mathbf{A}$. For complex matrices $\mathbf{A}$, $\mathbf{B}$ of appropriate size, one can take the conjugate transpose of the product $\mathbf{AB}$. If we denote the matrix full of conjugates of entries found in $\mathbf{A}$ by $\overline{\mathbf{A}}$, then we have

$$(\mathbf{AB})^{\mathrm{H}} = (\overline{\mathbf{AB}})^{\mathrm{T}} = (\overline{\mathbf{A}}\,\overline{\mathbf{B}})^{\mathrm{T}} = \overline{\mathbf{B}}^{\mathrm{T}}\overline{\mathbf{A}}^{\mathrm{T}} = \mathbf{B}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}.$$

In particular, the conjugate transpose of a matrix-vector product $\mathbf{Av}$ is

$$(\mathbf{Av})^{\mathrm{H}} = \mathbf{v}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}},$$

and if we need to take an inner product between $\mathbf{Au}$ and $\mathbf{v}$, we have the convenient formula

$$\langle \mathbf{Au}, \mathbf{v} \rangle \;=\; (\mathbf{Au})^{\mathrm{H}}\mathbf{v} \;=\; \mathbf{u}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}\mathbf{v} \;=\; \langle \mathbf{u}, \mathbf{A}^{\mathrm{H}}\mathbf{v} \rangle. \tag{2}$$

Formula (2) holds whether or not $\mathbf{A}$ is square.

Note that $\mathbf{A}^{\mathrm{H}}$ is sometimes called the **adjoint** matrix (as opposed to the *conjugate transpose*). If the entries in $\mathbf{A}$ are real numbers only, then $\mathbf{A}^{\mathrm{H}} = \mathbf{A}^{\mathrm{T}}$. Any matrix which satisfies $\mathbf{A}^{\mathrm{H}} = \mathbf{A}$ (necessarily square) is said to be **self-adjoint**, or **Hermitian**.

Symmetric Matrices

We use the term **symmetric** to describe a matrix $\mathbf{A}$ which satisfies $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$. A complex symmetric matrix need not be self-adjoint, but a real symmetric matrix is. Symmetric real matrices are very important in applications. They have some very favorable properties, some of which hold for all self-adjoint matrices. Every fact we state for self-adjoint matrices is true of symmetric real matrices.

Theorem 1: Eigenvalues of a self-adjoint matrix are real.

Proof: To prove this, we note first that any complex number z can be expressed in the form $z = a + ib$; here a, b are real numbers, called the *real* and *imaginary* parts of z, respectively. The number z is, in fact, *real* precisely when its imaginary part $b = 0$. Furthermore, the difference of z and its conjugate is

$$z - \bar{z} \;=\; (a + ib) - (a - ib) \;=\; i(2b),$$

which is zero if and only if $z \in \mathbb{R}$.

Now, suppose $(\lambda, \mathbf{v})$ is an eigenpair (with $\mathbf{v} \neq \mathbf{0}$) of a self-adjoint matrix $\mathbf{A}$. Consider the quantity $\lambda\|\mathbf{v}\|^2$, which may alternatively be expressed as

$$\lambda \langle \mathbf{v}, \mathbf{v} \rangle \;=\; \langle \mathbf{v}, \lambda\mathbf{v} \rangle \;=\; \langle \mathbf{v}, \mathbf{Av} \rangle \;=\; \langle \mathbf{A}^{\mathrm{H}}\mathbf{v}, \mathbf{v} \rangle \;=\; \langle \mathbf{Av}, \mathbf{v} \rangle \;=\; \langle \lambda\mathbf{v}, \mathbf{v} \rangle \;=\; \overline{\lambda} \langle \mathbf{v}, \mathbf{v} \rangle.$$

Subtracting the expression at one end from that on the other gives

$$0 \;=\; \lambda \langle \mathbf{v}, \mathbf{v} \rangle - \overline{\lambda} \langle \mathbf{v}, \mathbf{v} \rangle \;=\; (\lambda - \overline{\lambda}) \langle \mathbf{v}, \mathbf{v} \rangle \;=\; (\lambda - \overline{\lambda})\|\mathbf{v}\|^2.$$

Since $\|\mathbf{v}\| \neq 0$, it follows that $\lambda - \overline{\lambda} = 0$, which implies $\lambda \in \mathbb{R}$. □

As well, self-adjoint matrices generate eigenvectors which are naturally orthogonal.

Theorem 2: Eigenvectors corresponding to distinct eigenvalues of a self-adjoint matrix are orthogonal.

Proof: Suppose $(\mu, \mathbf{u})$, $(\lambda, \mathbf{v})$ are both eigenpairs of a self-adjoint matrix $\mathbf{A}$ with $\mu \neq \lambda$. By the previous theorem, μ and λ are real numbers, so $\overline{\mu} = \mu$. We have

$$\begin{aligned}(\mu - \lambda)\langle \mathbf{u}, \mathbf{v}\rangle &= \mu\langle \mathbf{u}, \mathbf{v}\rangle - \lambda\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mu\mathbf{u}, \mathbf{v}\rangle - \langle \mathbf{u}, \lambda\mathbf{v}\rangle = \langle \mathbf{A}\mathbf{u}, \mathbf{v}\rangle - \langle \mathbf{u}, \mathbf{A}\mathbf{v}\rangle \\ &= \left\langle \mathbf{u}, \mathbf{A}^{\mathrm{H}}\mathbf{v}\right\rangle - \langle \mathbf{u}, \mathbf{A}\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{A}\mathbf{v}\rangle - \langle \mathbf{u}, \mathbf{A}\mathbf{v}\rangle = 0.\end{aligned}$$

Since $(\mu - \lambda)\langle \mathbf{u}, \mathbf{v}\rangle = 0$ and $\mu - \lambda \neq 0$, it follows that $\langle \mathbf{u}, \mathbf{v}\rangle = 0$. □

Now, if $\mathbf{A}$ is self-adjoint (real and symmetric, for instance) and has n *distinct* eigenvalues (all real, of course) $\lambda_1, \ldots, \lambda_n$, then

- the corresponding eigenspaces Null$(\mathbf{A} - \lambda_j\mathbf{I})$ are all 1-dimensional (since GM = AM = 1 for each eigenvalue) having a single basis eigenvector $\mathbf{v}_j$,
- the vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ are linearly independent (proved earlier), and form a basis of $\mathbb{C}^n$, and
- the basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is *orthogonal*, not requiring a Gram-Schmidt process to make it so.

If we choose (or make) the lengths of the eigenvalues be 1, say, setting

$$\mathbf{q}_j = \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|}, \qquad j = 1, 2, \ldots, n,$$

then the square matrix $\mathbf{Q}$ having these vectors as its columns

$$\mathbf{Q} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix},$$

satisfies the relationship $\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{I}$. Such a $\mathbf{Q}$ is called a **unitary matrix**. Note that $\mathbf{Q}^{-1} = \mathbf{Q}^{\mathrm{H}}$. Naturally, $\mathbf{Q}$ serves to diagonalize $\mathbf{A}$, with

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathrm{H}}.$$

Note if $\mathbf{A}$ is real and symmetric, then the entries of $\mathbf{Q}$ can be taken to be real as well, making $\mathbf{Q}$ an orthogonal matrix whose columns, simple eigenvectors of $\mathbf{A}$, form an orthonormal basis of $\mathbb{R}^n$. In that case, the above can be written as

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathrm{T}}.$$

Now, if degenerate matrices (those with at least one eigenvalue for which GM $<$ AM) are undesirable, here is the really great news. Even though a self-adjoint matrix can have repeated eigenvalues—eigenvalues with AM > 1—no eigenvalue will have GM $<$ AM. This fact, which we do not prove at this time, is worthy of a gray box.

Fact 1: If μ is an eigenvalue of a self-adjoint matrix $\mathbf{A}$, then the algebraic multiplicity of μ (i.e., the maximum power m for which $(\lambda - \mu)^m$ is a factor of the characteristic polynomial $\det(\mathbf{A} - \lambda\mathbf{I})$) is equal to its geometric multiplicity (i.e., the dimension of the eigenspace $E_\mu = \text{Null}(\mathbf{A} - \mu\mathbf{I})$).

For a given eigenvalue λ whose AM $= k > 1$, it is still true that $\text{Null}(\mathbf{A} - \lambda\mathbf{I})$ has many bases. But, using Gram-Schmidt, it is always possible to choose an orthonormal basis $\{\mathbf{q}_1, \ldots, \mathbf{q}_k\}$ of eigenvectors for $\text{Null}(\mathbf{A} - \lambda\mathbf{I})$. This set of vectors, put together with orthonormal bases of the other eigenspaces, generate one complete orthonormal basis of $\mathbb{R}^n$. These various results are summarized in the Spectral Theorem.

Theorem 3 (Spectral Theorem for Self-Adjoint Matrices): Suppose $\mathbf{A}$ is a self-adjoint matrix. Then there exists an orthonormal basis of $\mathbb{C}^n$ consisting of eigenvectors $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ of $\mathbf{A}$. $\mathbf{A}$ is, hence, diagonalizable, and the matrix $\mathbf{Q}$ whose j^{th} column is $\mathbf{q}_j$, $j = 1, \ldots, n$, is a unitary matrix which serves to diagonalize $\mathbf{A}$, so that $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathrm{T}}$. As a special case, if $\mathbf{A}$ is real and symmetric, then there exists an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ of $\mathbf{A}$. The matrix $\mathbf{Q}$ formed from these eigenvectors serves to diagonalize $\mathbf{A}$, so that $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathrm{T}}$.

The result of the spectral theorem can be realized in steps like those outlined in the following example.

Example 1:

Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \sqrt{2} \\ 0 & 2 & 0 \\ \sqrt{2} & 0 & 0 \end{bmatrix}.$$

Since $\mathbf{A}$ is a symmetric matrix, the Spectral Theorem guarantees $\mathbb{R}^3$ has a basis which consists of eigenvectors of $\mathbf{A}$. The following steps lead to a realization of such a basis.

1. **Find the eigenvalues**. The process is the same as for finding eigenvalues of any square matrix, except we know the results will be real numbers. They are, in fact $\lambda = -1, 2, 2$ (i.e., 2 is an eigenvalue with AM = 2).

2. **Find bases of the various eigenspaces**. The eigenspace E_{-1} consists of solutions to $(\mathbf{A} + \mathbf{I})\mathbf{v} = \mathbf{0}$. The augmented matrix

$$\left[(\mathbf{A}+\mathbf{I})\,\middle|\,\mathbf{0}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 1/\sqrt{2} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

has, as expected, one free column, leading to eigenvector $\mathbf{v}_1 = (-1, 0, \sqrt{2})$.

The eigenspace E_2 consists of solutions to $(\mathbf{A} - 2\mathbf{I})\mathbf{v} = \mathbf{0}$. The augmented matrix

$$\left[(\mathbf{A}-2\mathbf{I})\,\middle|\,\mathbf{0}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -\sqrt{2} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

has two free columns, matching the algebraic multiplicity of the eigenvalue 2. One can verify that $\mathbf{v}_2 = (\sqrt{2}, 0, 1)$ and $\mathbf{v}_3 = (0, 1, 0)$ are independent eigenvectors in E_2.

3. **For those eigenspaces of dimension > 1, find orthogonal bases using Gram-Schmidt**. In this instance, E_2 is the only eigenspace of dimension higher than 1. As luck would have it, the basis $\{\mathbf{v}_2, \mathbf{v}_3\}$ is already orthogonal.

4. **Amass the bases of the various eigenspaces and normalize**. It is already the case that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis of $\mathbb{R}^3$ consisting of eigenvectors of $\mathbf{A}$. The steps prior to this one ensure they form, in fact, an orthogonal basis. We now set

$$\begin{aligned}
\mathbf{q}_1 &= \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \frac{1}{\sqrt{3}}\begin{bmatrix} -1 \\ 0 \\ \sqrt{2} \end{bmatrix} = \begin{bmatrix} -1/\sqrt{3} \\ 0 \\ \sqrt{2/3} \end{bmatrix} \\
\mathbf{q}_2 &= \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \frac{1}{\sqrt{3}}\begin{bmatrix} \sqrt{2} \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \sqrt{2/3} \\ 0 \\ 1/\sqrt{3} \end{bmatrix} \\
\mathbf{q}_3 &= \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.
\end{aligned}$$

Then

$$\mathbf{Q} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix} = \begin{bmatrix} -1/\sqrt{3} & \sqrt{2/3} & 0 \\ 0 & 0 & 1 \\ \sqrt{2/3} & 1/\sqrt{3} & 0 \end{bmatrix}$$

is an orthogonal matrix, and

$$\mathbf{A} = \mathbf{Q}\begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}\mathbf{Q}^{\mathrm{T}} = -\mathbf{q}_1\mathbf{q}_1^{\mathrm{T}} + 2\mathbf{q}_2\mathbf{q}_2^{\mathrm{T}} + 2\mathbf{q}_3\mathbf{q}_3^{\mathrm{T}}.$$

■