A.1 Complex Arithmetic

A *complex number* is a quantity of the form $x + yi$, where x and y are real numbers and $i^2 = -1$. Calculations with complex numbers are not difficult. Just use all of the usual algebraic rules, and replace i^2 with -1 whenever you can. For example, suppose that $\chi = x+yi$ is one complex number and $\omega = z+wi$ is another. Here's addition:

$$\chi + \omega = (x + yi) + (z + wi) = (x + z) + (y + w)i.$$

Similarly, subtraction is

$$\chi - \omega = (x + yi) - (z + wi) = (x - z) + (y - w)i.$$

Multiplication is a little more interesting:

$$\chi\omega = (x + yi)(z + wi) = xz + xwi + yiz + ywi^2 = (xz - yw) + (xw + yz)i.$$

Division is more difficult. It's helpful to think of division as multiplication by the reciprocal:

$$\frac{\chi}{\omega} = \frac{x + yi}{z + wi} = (x + yi) \cdot \frac{1}{z + wi} = \chi\omega^{-1}.$$

But how do you compute the reciprocal? Use this trick:

$$\omega^{-1} = \frac{1}{z + wi} = \frac{1}{z + wi} \cdot \frac{z - wi}{z - wi} = \frac{z - wi}{z^2 + w^2} = \frac{z}{z^2 + w^2} + \frac{-w}{z^2 + w^2}i. \tag{A.1}$$

It doesn't work if $z^2 + w^2 = 0$, but that happens only when you're trying to compute the reciprocal of $\omega = 0 + 0i = 0$. Division by zero is illegal in the complex numbers, just as it's illegal in the real numbers.

Exercise A.1.1. *Here is some practice with complex arithmetic. If you want more practice, then make up your own problems.*

1. *Compute* $(1 + i)^4$.
2. *Compute* $(2 + i)/(3 - 4i)$.
3. *Using the quadratic formula (or another method), solve* $\chi^2 + 2\chi + 2 = 0$.

The set of complex numbers is denoted $\mathbb{C}$. When I say that the complex numbers satisfy all of the usual rules of algebra, I specifically mean these nine rules:

- Associativity of addition: $(\chi + \omega) + \psi = \chi + (\omega + \psi)$ for all $\chi, \omega, \psi \in \mathbb{C}$.
- Identity in addition: The complex number $0 = 0 + 0i$ satisfies $0 + \chi = \chi = \chi + 0$ for all χ.
- Inverses in addition: For any complex number χ, there exists a complex number $-\chi$, which satisfies $\chi + -\chi = 0 = -\chi + \chi$. (Explicitly, if $\chi = x + yi$, then $-\chi = -x + -yi$.)
- Commutativity of addition: $\chi + \omega = \omega + \chi$ for all χ, ω.
- Associativity of multiplication: $(\chi\omega)\psi = \chi(\omega\psi)$ for all χ, ω, ψ.
- Identity in multiplication: The complex number $1 = 1 + 0i$ satisfies $1\chi = \chi = \chi 1$ for all χ.

- Inverses in multiplication: For any $\chi \neq 0$, there exists a complex number χ^{-1}, which satisfies $\chi\chi^{-1} = 1 = \chi^{-1}\chi$. (Equation A.1 tells us how to compute χ^{-1}.)
- Commutativity of multiplication: $\chi\omega = \omega\chi$ for all χ, ω.
- Distributivity: $\chi(\omega + \psi) = \chi\omega + \chi\psi$ and $(\chi + \omega)\psi = \chi\psi + \omega\psi$ for all χ, ω, ψ.

In the mathematical jargon, we say that $\mathbb{C}$ is a *field.* If you don't enjoy thinking about abstract mathematical structures defined by axioms, that's fine. Just think of the list above as a crib sheet of the basic rules that complex numbers obey.

The set $\mathbb{R}$ of real numbers is also a field. In fact, $\mathbb{R}$ can be viewed as a subset of $\mathbb{C}$, in that the real number x can be identified with the complex number $x + 0i$. In one way, $\mathbb{R}$ is nicer than $\mathbb{C}$: $\mathbb{R}$ has an ordering $<$, so that we can talk about whether $x < z$, $x \geq z$, etc. for real numbers x, z. Those concepts don't exist in $\mathbb{C}$. But $\mathbb{C}$ is nicer than $\mathbb{R}$ in a different way: It is *algebraically closed*, meaning that every non-constant polynomial with coefficients in $\mathbb{C}$ has a root in $\mathbb{C}$. In contrast, there are polynomials with real coefficients that do not have any real roots. Arguably the most important example is $x^2 + 1$. Why?

The complex numbers have another operation, which has no analogue in the real numbers: conjugation. The *conjugate* of a complex number $\chi = x + yi$ is defined as

$$\overline{\chi} = \overline{x + yi} = x - yi.$$

Notice that the conjugate of a real number $x = x + 0i$ is just $x - 0i = x$ again. So conjugation of real numbers is trivial, which is why we never talk about it. Notice also that

$$\chi\overline{\chi} = (x + yi) \cdot \overline{x + yi} = x^2 + y^2.$$

This trick helped us compute the reciprocal in Equation A.1. Conjugation also plays well with arithmetic, as the following exercise shows.

Exercise A.1.2. *Prove that* $\overline{\chi + \omega} = \overline{\chi} + \overline{\omega}$ *and* $\overline{\chi\omega} = \overline{\chi}\,\overline{\omega}$ *for all* $\chi, \omega \in \mathbb{C}$.

A.2 Complex Geometry

Because each complex number is made up of two real numbers, it is natural to picture $\mathbb{C}$ as the two-dimensional real plane $\mathbb{R}^2$. That is, the number $\chi = x + yi$ plots at the point (x, y). The horizontal axis consists of the numbers of the form $x + 0i$ — that is, the real numbers. The vertical axis consists of the numbers of the form $0 + yi$. They are called the *imaginary* numbers.

The terms "real" and "imaginary" are important in the vocabulary of mathematics, so you should learn to use them correctly. First, they are not antonyms. Most complex numbers are neither real nor imaginary, and the complex number $0 = 0 + 0i$ is both real and imaginary. If you want to say that a number is not real, then don't say that it's imaginary; instead, say that it's "not real" or "non-real". Second, you should ignore the non-mathematical meanings of "real" and "imaginary". You should not intuit that the real numbers actually exist and the other complex numbers actually don't exist. None of these numbers exist in our universe. They are concepts, not physical objects, and they live only in the human mind.

Once we view complex numbers as points or vectors in $\mathbb{R}^2$, complex addition has a simple geometric interpretation. Adding a complex number $\chi = x + yi$ to a complex number $\omega = z + wi$ has the effect of translating the point ω by x units to the right and y units up. Scaling $\omega = z+wi$ by a real $\chi = x+0i = x$ has the effect of stretching ω away from the origin by a factor of x. The *norm* or *magnitude* of a complex number $\chi = x + yi$ is defined as its distance to the origin:

$$|\chi| = |x + yi| = \sqrt{x^2 + y^2} = \sqrt{(x + yi) \cdot \overline{x + yi}} = \sqrt{\chi\overline{\chi}}.$$

In other words, when we view $\mathbb{C}$ as the vector space $\mathbb{R}^2$, then addition, real scaling, and the norm have their usual geometric interpretation. (In the jargon of mathematics, $\mathbb{C}$ and $\mathbb{R}^2$ are isomorphic as normed two-dimensional vector spaces over $\mathbb{R}$.)

Exercise A.2.1. *Let $\chi, \omega \in \mathbb{C}$ be arbitrary.*

1. *Prove that $\chi + \overline{\chi}$ is real and $\chi + \overline{\chi} \leq 2|\chi|$.*
2. *Prove the triangle inequality: $|\chi + \omega| \leq |\chi| + |\omega|$.*

However, $\mathbb{C}$ has two more features that $\mathbb{R}^2$ lacks: conjugation and general complex multiplication. Geometrically, conjugation has the effect of flipping points across the real axis. To understand the geometric meaning of multiplication, it is helpful to change coordinates.

Recall (from some calculus course) the concept of polar coordinates. Given a point (x, y) in the plane, let r be the distance from the origin to that point, and let t be the angle, at the origin, measured counterclockwise from the positive real axis to (x, y). It is easy to convert from polar coordinates (r, t) to Cartesian coordinates (x, y): $x = r \cos t$ and $y = r \sin t$. So

$$x + iy = r \cos t + ir \sin t = r(\cos t + i \sin t).$$

To convert from Cartesian to polar coordinates, set $r = \sqrt{x^2 + y^2}$ and

$$t = \begin{cases} \arctan{(y/x)} & \text{if } x > 0, \\ \arctan{(y/x)} + \pi & \text{if } x < 0, \\ \pi/2 & \text{if } x = 0 \text{ and } y > 0, \\ -\pi/2 & \text{if } x = 0 \text{ and } y < 0. \end{cases} .$$

Most programming languages offer a function, called something like `atan2`, to compute t conveniently and robustly. (When $x = 0$ and $y = 0$, $r = 0$ and t is undefined. Often t is arbitrarily declared to be 0, with no practical harm.)

Suppose that we have two complex numbers expressed in polar coordinates:

$$\chi = r(\cos t + i \sin t) \text{ and } \omega = s(\cos u + i \sin u).$$

Then, using a couple of trigonometric identities, we can compute

$$\begin{aligned} \chi\omega &= r(\cos t + i \sin t)\ s(\cos u + i \sin u) \\ &= (rs)\left((\cos t \cos u - \sin t \sin u) + i(\cos t \sin u + \sin t \cos u)\right) \\ &= (rs)\left(\cos(t + u) + i \sin(t + u)\right). \end{aligned}$$

The norm of this complex number is rs, and the angular coordinate is $t+u$. So multiplication of complex numbers amounts to multiplying their norms and adding their angles. In other words, multiplying ω by $r(\cos t + i \sin t)$ scales ω away from the origin by a factor of r and rotates ω through an angle of t about the origin.

Exercise A.2.2. *We have given geometric interpretations for addition, multiplication, and conjugation. For example, conjugation flips $\mathbb{C}$ across $\mathbb{R}$. Now what is the geometric interpretation of inversion? I mean, consider the map $f : (\mathbb{C} - \{0\}) \to \mathbb{C}$ given by $f(\chi) = \chi^{-1}$. In English and maybe pictures, describe the geometric effect of this map.*

A.3 Complex Exponentiation

In the real numbers, the exponential function is defined by the power series

$$\exp(x) = \sum_{k=0}^{\infty} \frac{1}{k!} x^k = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4 + \cdots . \tag{A.2}$$

This function has many miraculous properties, the most important of which is probably

$$\exp(x) \cdot \exp(y) = \exp(x+y).$$

Let's plug our favorite real numbers into exp. First, $\exp(0) = 1$. Second,

$$\exp(1) = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots = 2.718\ldots.$$

We give the number $\exp(1)$ a special name: e. Then $\exp(2) = \exp(1+1) = \exp(1) \cdot \exp(1) = e^2$. Using induction, you can prove that $\exp(n) = e^n$ for any positive integer n, and then for all integers n. For this reason, the function $\exp(x)$ is often denoted e^x. (But the function exp is more fundamental than the number e. You should view the number as an emergent phenomenon of the function.)

The same power series function definition (Equation A.2) works for complex numbers χ. I mean, you can plug in any complex number $\chi = x + yi$ for x, compute the required powers, divide by the required factorials, and perform the required summation (at least in principle). The complex exponential function still has that crucial sum-product property

$$\exp(\chi) \cdot \exp(\omega) = \exp(\chi + \omega).$$

You already know what exp does to complex numbers χ of the form $x + 0i$, because those are just real numbers. But what about imaginary numbers $\chi = 0 + iy$? When one examines the power series closely, something surprising happens: $e^{iy} = \cos y + i \sin y$. The exponential function contains the trigonometric functions, even though Equation A.2 seems not to be related to trigonometry at all.

Exercise A.3.1. *Using the power series*

$$\cos y = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} \text{ and } \sin y = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1},$$

prove the claim above that $e^{iy} = \cos y + i \sin y$. (At some point you need to rearrange the terms in a series. Depending on your training, you might know that rearranging is dangerous. But stop worrying; these are the nicest series in the world.)

The exponential function gives us another way to view polar coordinates:

$$\chi = x + iy = r(\cos t + i \sin t) = re^{it}.$$

Multiplication is especially simple in this format. If $\omega = z + wi = se^{iu}$ is another complex number, then

$$\chi\omega = (a + bi)(c + di) = re^{it}se^{iu} = (rs)e^{i(t+u)}.$$

Exercise A.3.2. *This exercise explores the geometry of the exponential map $\chi \mapsto \exp(\chi)$.*

1. *As χ varies over the real numbers, $\exp(\chi)$ varies over what subset of $\mathbb{C}$?*

2. *As χ varies over the imaginary numbers, $\exp(\chi)$ varies over what subset of $\mathbb{C}$?*

3. *For which χ is $\exp(\chi)$ real? The set of such χ forms a subset of $\mathbb{C}$; describe it.*

4. *For which χ is $\exp(\chi)$ imaginary?*

Exercise A.3.3. *Prove, for any real number t, that $\left|e^{i2t} - 1\right|^2 = 4 \sin^2 t$. (Hint: You might want to use a double-angle trigonometric identity.)*

A.4 Complex Matrices

Your linear algebra course was probably focused on vector spaces over the real numbers, meaning that all scalars were real numbers. But linear algebra works just as well over the complex numbers. In some ways, linear algebra works better over $\mathbb{C}$ than over $\mathbb{R}$, actually.

A $p \times q$ *complex matrix* is a rectangular array of complex numbers with p rows and q columns. Because we are doing computer science, we index the rows and columns from 0 rather than from 1. That is, the rows are numbered $0, \ldots, p-1$ instead of $1, \ldots, p$, and the columns are numbered $0, \ldots, q-1$ instead of $1, \ldots, q$.

Two $p \times q$ matrices can be added component-wise:

$$\begin{aligned} A + B &= \begin{bmatrix} A_{00} & \cdots & A_{0,q-1} \\ \vdots & \ddots & \vdots \\ A_{p-1,0} & \cdots & A_{p-1,q-1} \end{bmatrix} + \begin{bmatrix} B_{00} & \cdots & B_{0,q-1} \\ \vdots & \ddots & \vdots \\ B_{p-1,0} & \cdots & B_{p-1,q-1} \end{bmatrix} \\ &= \begin{bmatrix} A_{00} + B_{00} & \cdots & A_{0,q-1} + B_{0,q-1} \\ \vdots & \ddots & \vdots \\ A_{p-1,0} + B_{p-1,0} & \cdots & A_{p-1,q-1} + B_{p-1,q-1} \end{bmatrix}. \end{aligned}$$

And they can be scaled by scalars $\sigma \in \mathbb{C}$:

$$\sigma A = \sigma \begin{bmatrix} A_{00} & \cdots & A_{0,q-1} \\ \vdots & \ddots & \vdots \\ A_{p-1,0} & \cdots & A_{p-1,q-1} \end{bmatrix} = \begin{bmatrix} \sigma A_{00} & \cdots & \sigma A_{0,q-1} \\ \vdots & \ddots & \vdots \\ \sigma A_{p-1,0} & \cdots & \sigma A_{p-1,q-1} \end{bmatrix}.$$

It's all very much like real linear algebra, except that the underlying scalar additions and multiplications are of complex numbers rather than real numbers. Consequently, addition and scalar multiplication of $p \times q$ matrices obey all of the algebraic rules that you'd expect. The most basic of these rules are called the axioms for a *vector space*.

- Associativity of addition: $(A+B)+C = A+(B+C)$.
- Identity in addition: There exists a zero matrix 0 such that $0+A = A = A+0$.
- Inverses in addition: For all A, there exists a $-A$, which satisfies $A + -A = 0 = -A + A$.
- Commutativity of addition: $A+B = B+A$.
- Associativity of multiplication: $(\sigma\tau)A = \sigma(\tau A)$.
- Identity in multiplication: The complex number 1 satisfies $1A = A$.
- Distributivity: $\sigma(A+B) = \sigma A + \sigma B$ and $(\sigma+\tau)A = \sigma A + \tau B$.

A $p \times q$ matrix A can be multiplied by a $q \times m$ matrix B just as you'd expect:

$$\begin{aligned} A \cdot B &= \begin{bmatrix} A_{00} & \cdots & A_{0,q-1} \\ \vdots & \ddots & \vdots \\ A_{p-1,0} & \cdots & A_{p-1,q-1} \end{bmatrix} \cdot \begin{bmatrix} B_{00} & \cdots & B_{0,m-1} \\ \vdots & \ddots & \vdots \\ B_{q-1,0} & \cdots & B_{q-1,m-1} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{k=0}^{q-1} A_{0k}B_{k0} & \cdots & \sum_{k=0}^{q-1} A_{0,k}B_{k,q-1} \\ \vdots & \ddots & \vdots \\ \sum_{k=0}^{q-1} A_{p-1,k}B_{k,0} & \cdots & \sum_{k=0}^{q-1} A_{p-1,k}B_{k,q-1} \end{bmatrix}. \end{aligned}$$

In particular, $p \times p$ complex matrices can be multiplied with each other. Under the operations of addition and matrix multiplication, the set of $p \times p$ complex matrices forms a *non-commutative ring with identity* — meaning, it satisfies all of the field axioms (Section A.1) except for two: commutativity of multiplication and inverses in multiplication. For example, $A(B+C) = AB + AC$, and the identity matrix I satisfies $IA = A = AI$. Additionally the set of $p \times p$ complex matrices satisfies the following property, which together with the vector space axioms and ring axioms makes it an *associative algebra*.

- $\sigma(AB) = (\sigma A)B = A(\sigma B)$.

The *transpose* of a $p \times q$ matrix A is the $q \times p$ matrix obtained by reflecting the entries across the main diagonal:

$$A^\top = \begin{bmatrix} A_{00} & \cdots & A_{0,q-1} \\ \vdots & \ddots & \vdots \\ A_{p-1,0} & \cdots & A_{p-1,q-1} \end{bmatrix}^\top = \begin{bmatrix} A_{00} & \cdots & A_{p-1,0} \\ \vdots & \ddots & \vdots \\ A_{0,q-1} & \cdots & A_{p-1,q-1} \end{bmatrix}.$$

Just as in real linear algebra, transposition interacts with arithmetic as follows.

$$\begin{aligned} (A+B)^\top &= A^\top + B^\top, \\ (\sigma A)^\top &= \sigma A^\top, \\ (AB)^\top &= B^\top A^\top. \end{aligned}$$

What's new in complex matrices, compared to real matrices, is the operation of conjugation. It's really easy; you just conjugate each entry:

$$\overline{A} = \overline{\begin{bmatrix} A_{00} & \cdots & A_{0,q-1} \\ \vdots & \ddots & \vdots \\ A_{p-1,0} & \cdots & A_{p-1,q-1} \end{bmatrix}} = \begin{bmatrix} \overline{A_{00}} & \cdots & \overline{A_{0,q-1}} \\ \vdots & \ddots & \vdots \\ \overline{A_{p-1,0}} & \cdots & \overline{A_{p-1,q-1}} \end{bmatrix}.$$

Because conjugation respects addition and multiplication, it also respects all three operations of matrix arithmetic:

$$\begin{aligned}\overline{A+B} &= \overline{A}+\overline{B},\\ \overline{\sigma A} &= \overline{\sigma}\overline{A},\\ \overline{AB} &= \overline{A}\,\overline{B}.\end{aligned}$$

Sometimes it happens that a $p \times p$ matrix A, a non-zero $p \times 1$ matrix B, and a scalar λ satisfy the equation $AB = \lambda B$. In this case, we say that B is an *eigenvector* for A with *eigenvalue* λ. It follows that σB is also an eigenvector with the same eigenvalue, for any scalar σ.

A nice feature of complex linear algebra, compared to real linear algebra, is that a $p \times p$ matrix A always has exactly p eigenvalues. (They might be identical, and they might share the same eigenvector. For example, for any nonzero $\chi \in \mathbb{C}$,

$$\left[\begin{array}{cc} 1 & \chi \\ 0 & 1 \end{array}\right]$$

has eigenvalues 1 and 1, both with eigenvector $B = [1\ \ 0]^\top$.) The *trace* of A is the sum of the eigenvalues of A. It can be computed easily as the sum of the diagonal entries:

$$\operatorname{tr} A = \sum_{j=1}^{p} A_{jj}.$$

The determinant of A is the product of the eigenvalues. The determinant is computed just as in real linear algebra. For example, the 2×2 case is

$$\det A = \det \left[\begin{array}{cc} A_{00} & A_{01} \\ A_{10} & A_{11} \end{array}\right] = A_{00}A_{11} - A_{01}A_{10}.$$

Just as in real linear algebra, $\det(AB) = (\det A)(\det B)$, and A^{-1} exists if and only if $\det A \neq 0$.

A.5 Complex Inner Product

In real linear algebra, the dot product is another fundamental operation. All of Euclidean geometry is a consequence of the dot product. For example, the length of a vector $\vec{v}$ is $\sqrt{\vec{v} \cdot \vec{v}}$, and the angle between two unit vectors $\vec{v}$ and $\vec{w}$ is $\arccos(\vec{v} \cdot \vec{w})$. The dot product relates to transposition via the equation $\vec{v} \cdot \vec{w} = \vec{v}^\top \vec{w}$. In fact, the entire transposition operation on matrices exists because of the dot product.

The best way to extend these concepts to complex matrices is to replace transposition with conjugate transposition. Define the *conjugate-transpose* A^* as

$$A^* = \overline{A}^\top = \overline{A^\top}.$$

It behaves much like transposition with respect to matrix arithmetic, but notice the conjugation that creeps in:

$$\begin{aligned}(A+B)^* &= A^* + B^*,\\ (\sigma A)^* &= \overline{\sigma} A^*,\\ (AB)^* &= B^* A^*.\end{aligned}$$

For defining the complex analogue of the dot product, there are a couple of conventions. We follow the convention used in the quantum computing textbooks of Mermin [2007], **?**. If A and B are $p \times q$ complex matrices, then their *Frobenius inner product* is

$$\langle A|B\rangle = \operatorname{tr}\left(A^*B\right) = \sum_{j=1}^{q} \left(A^*B\right)_{jj} = \sum_{j=1}^{q}\sum_{k=1}^{p} \overline{A_{kj}} B_{kj}.$$

Consider the special case where $q = 1$, so that A and B are p-dimensional column vectors. Then A^*B is 1×1, so we can regard it as a complex number, and we can write the inner product more simply: $\langle A|B\rangle = A^*B$. This operation on column vectors is called the *Hermitian inner product.*

The Frobenius inner product (including the special case of the Hermitian inner product) satisfies the following rules.

- Linearity in the second argument: $\langle A|\sigma B + \tau C\rangle = \sigma\langle A|B\rangle + \tau\langle A|C\rangle$.
- Conjugate symmetry: $\langle A|B\rangle = \overline{\langle B|A\rangle}$.
- Conjugate linearity in the first argument: $\langle \sigma A + \tau B|C\rangle = \overline{\sigma}\langle A|C\rangle + \overline{\tau}\langle B|C\rangle$.
- Positive definiteness: $\langle A|A\rangle$ is a positive real number, except when A is the zero matrix, in which case $\langle A|A\rangle = 0$.

Exercise A.5.1. *Prove the third property from the first two.*

With the inner product in hand, now we can define the *norm* or *magnitude* of any $p \times q$ matrix A to be the real number

$$\|A\| = \sqrt{\langle A|A\rangle} = \left(\sum_{j=1}^{q}\sum_{k=1}^{p} \overline{A_{kj}} A_{kj}\right)^{1/2} = \left(\sum_{j=1}^{q}\sum_{k=1}^{p} \left|A_{kj}\right|^2\right)^{1/2}.$$

The norm satisfies the following rules.

- Positivity: $\|A\| > 0$, unless A is the zero matrix, in which case $\|A\| = 0$.
- Scaling: $\|\sigma A\| = |\sigma| \cdot \|A\|$.
- Triangle inequality: $\|A + B\| \leq \|A\| + \|B\|$.
- Cauchy-Schwarz inequality: $|\langle A|B\rangle|^2 \leq \langle A|A\rangle \cdot \langle B|B\rangle$.

A $p \times 1$ column vector of norm 1 is said to be a *unit* vector.

Exercise A.5.2. *Let A and B be arbitrary $p \times q$ matrices.*

1. *Prove the Cauchy-Schwarz inequality. (Hint: If A is the zero matrix, then check that the inequality holds. If A is not zero, then let $C = B - (\langle A|B\rangle/\langle A|A\rangle)\, A$, and use the fact that $\langle C|C\rangle \geq 0$.)*
2. *Use Cauchy-Schwarz to prove the triangle inequality.*

A.6 Unitary and Hermitian Matrices

A $p \times p$ matrix U is *unitary* if $UU^* = I = U^*U$. The set of all $p \times p$ unitary matrices is denoted $\mathrm{U}(p)$. It does not form a vector space. Instead, $\mathrm{U}(p)$ forms a *group*, meaning that:

- Closure: If U and V are unitary, then so is the product UV.
- Identity: The identity matrix I is a unitary matrix.
- Inverses: If U is unitary, then U is invertible, and U^{-1} is also unitary.

A helpful (and non-obvious) fact about unitary matrices is that they diagonalize unitarily. That is, if U is $p \times p$ unitary, then there exists a $p \times p$ unitary V and a $p \times p$ diagonal unitary D such that $U = VDV^*$.

Exercise A.6.1. *Prove the claims above, that* $\mathrm{U}(p)$ *is a group and not a vector space.*

Exercise A.6.2. *Let U be a unitary 2×2 matrix. Let A, B be any 2×1 matrices.*

1. *Prove that the inner product of A with B equals the inner product of UA with UB.*
2. *Prove that the norm of UA equals the norm of A.*
3. *Conversely, prove that if V is 2×2 and the norm of VA equals the norm of A for all 2×1 matrices A, then V is unitary. (Hint: Consider three cases: A is a standard basis vector, A is a sum of two distinct standard basis vectors, and A is the sum of a standard basis vector and i times a distinct one.)*
4. *Prove that the eigenvalues of U have norm 1, and so does* $\det U$.
5. *Prove that a diagonal 2×2 complex matrix is unitary if and only if its diagonal entries are complex numbers of norm 1.*

A $p \times p$ matrix Q is *Hermitian* if $Q^* = Q$. A matrix S is *skew-Hermitian* if $S^* = -S$. The set of all $p \times p$ Hermitian matrices forms a real vector space of dimension p^2. So does the set of all $p \times p$ skew-Hermitian matrices.

Exercise A.6.3. *Just for the cases $p = 2$ and $p = 4$, check that the real vector space of $p \times p$ skew-Hermitian matrices is p^2-dimensional. That is, explain why choosing a $p \times p$ skew-Hermitian matrix is tantamount to choosing p^2 real numbers.*

Exercise A.6.4. *Prove that if S is skew-Hermitian, then iS is Hermitian. (The converse is also true. And Q is Hermitian if and only if iQ is skew-Hermitian. But I'm not asking you to prove those extra theorems.)*

Because unitary matrices are important in our course, we want to have some ways to manufacture them. One way is via the matrix exponential function. If A is an $p \times p$ complex (or real) matrix, define

$$\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k = I + A + \frac{1}{2}A^2 + \frac{1}{6}A^3 + \frac{1}{24}A^4 + \cdots .$$

Exercise A.6.5. *Prove that if S is skew-Hermitian, then* $\exp(S)$ *is unitary. [Hint: First show that* $(\exp(S))^* = \exp(-S)$*. To do that, you must perform some term-by-term manipulations on the power series. If you are well trained in power series, then such manipulations should make you nervous. But this is the nicest power series in the world, so stop worrying.]*

So, to make an $p \times p$ unitary matrix, one approach is: Choose p^2 real numbers, form a skew-Hermitian matrix S from them, and then exponentiate to get a unitary $U = \exp(S)$. For example, let

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Then $\{I, X, Y, Z\}$ is a basis for the real vector space of 2×2 Hermitian matrices, and

$$U = \exp(i(wI + xX + yY + zZ))$$

is a 2×2 unitary matrix, for any real numbers w, x, y, z.

A last useful fact for this section is the spectral theorem: If Q is $p \times p$ Hermitian matrix, then there exist p orthonormal eigenvectors of Q, and their associated eigenvalues are all real. (Orthonormality means that the eigenvectors have norm 1 and their inner products with each other are all 0.)

A.7 Calculus with Matrices

Let t be a real variable. Think of it as time. Suppose that $A = A(t)$ is a time-dependent complex matrix. You can think of it as a matrix, each entry of which is a function of t. You can also think of it as a curve of matrices parametrized by t. Anyway, define $\frac{d}{dt}A = A' = \dot{A}$ to be the component-wise derivative of A with respect to t. For example,

$$\frac{d}{dt} \begin{bmatrix} 1 + t^2 & \cos t \\ e^t & 0 \end{bmatrix} = \begin{bmatrix} 2t & -\sin t \\ e^t & 0 \end{bmatrix}.$$

This matrix derivative obeys some reasonable rules:

- Sum rule: $\frac{d}{dt}(A + B) = \dot{A} + \dot{B}$.
- Scaling rule: $\frac{d}{dt}(\sigma A) = \sigma \dot{A}$ for any constant scalar σ.
- Product rule: $\frac{d}{dt}(AB) = \dot{A}B + A\dot{B}$. (Be careful about the order, because matrix multiplication is not commutative.)
- Conjugate-transpose rule: $\frac{d}{dt}(A^*) = (\dot{A})^*$. That is, the derivative of the conjugate-transpose is the conjugate-transpose of the derivative.

Exercise A.7.1. *Prove that if* $U = U(t)$ *is unitary, then* $\dot{U}U^*$ *is skew-Hermitian.*

For the next exercise, let $U = U(t)$ be a time-dependent $p \times p$ unitary matrix. Let B be a constant $p \times 1$ column vector. Therefore $A = UB$ is a time-dependent $p \times 1$ matrix.

Exercise A.7.2. *In this situation, prove that there exists a Hermitian* Q *such that* $i\dot{A} = QA$*. [Hint: First show that* $\dot{A} = SA$ *for a certain skew-Hermitian* S*.]*