

Vectors and Matrices

Vectors

Vector: A collection of complex or real numbers, generally put in a column

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix} \text{ where each } v_i \in C ; \text{ Dimension of } \mathbf{v} \text{ is } N$$

Transpose of a Vector: Sometimes it is helpful to deal with a row version of a vector, which is called the transpose of the vector and is denoted with a superscript T:

$$\mathbf{v}^T = [v_1 \quad \cdots \quad v_N]$$

We can also use the following variation of this:

$$\mathbf{v} = [v_1 \quad \cdots \quad v_N]^T$$

which is often convenient for notational purposes to save vertical space.

Vector Addition: $\mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_N \end{bmatrix} \quad \mathbf{a} + \mathbf{b} = \begin{bmatrix} a_1 + b_1 \\ \vdots \\ a_N + b_N \end{bmatrix}$

Scalar: A real or complex number. If the vectors of interest are complex valued then the scalars are taken to be complex; if the vectors of interest are real valued then the scalars are taken to be real.

Multiply by a Scalar: $\mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix} \quad \alpha \mathbf{a} = \begin{bmatrix} \alpha a_1 \\ \vdots \\ \alpha a_N \end{bmatrix}$

Note: multiplying a vector by a scalar is viewed as changing its length. If the scalars are real: (i) multiplying by a scalar whose magnitude is greater than 1 increases the length; (ii) multiplying by a scalar whose magnitude is less than 1 decreases the length; (iii) if the scalar is negative, multiplying by it “flips” the vector to point in the opposite direction

Arithmetic Properties of Vectors: vector addition and scalar multiplication exhibit the following properties. This basically just says that the arithmetic structure of vectors behaves pretty much like the real numbers – the main difference is that you can't multiply a vector by a vector, but you have a different set of numbers (the scalars) that come into play for the multiplication aspects. Let \mathbf{x} , \mathbf{y} , and \mathbf{z} be vectors of the same dimension and let α and β be scalars; then the following properties hold:

- (1) Commutativity $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
 $\alpha\mathbf{x} = \mathbf{x}\alpha$

- (2) Associativity $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{y} + (\mathbf{x} + \mathbf{z})$
 $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$

- (3) Distributivity $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$
 $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$

- (4) Unity and Zero Scalar $1\mathbf{x} = \mathbf{x}$
 $1\mathbf{x} = \mathbf{0}$, where $\mathbf{0}$ is the zero vector of all zeros

Vector Space: Set of all vectors of some fixed dimension N , together with 2 operations defined on the set:

- (i) addition of vectors – gives another vector in the set
- (ii) multiply a vector by a scalar – gives another vector in the set

In other words: a set vectors qualifies to be a vector *space* if and only if the set is closed under addition and closed under scalar multiplication.

Note: this means that ANY linear combination of vectors in the space results in a vector in the space.

Example: Consider the vector space C^N = set of all complex N -Vectors with complex numbers as the set of scalars. Then, any linear combination of such vectors is in C^N . For example, let $\{\mathbf{v}_1, \dots, \mathbf{v}_K\}$ be any set of K vectors in C^N and let $\{\alpha_1, \dots, \alpha_K\}$ be any set of K scalars in C , then

$$\sum_{k=1}^K \alpha_k \mathbf{v}_k \in C^N$$

Axioms of Vector Space: If V is a set of vectors satisfying the above definition of a vector space then it satisfies the following axioms – the first four are simply the arithmetic properties of vectors:

- (1) Commutativity (see above)
- (2) Associativity (see above)
- (3) Distributivity (see above)
- (4) Unity and Zero Scalar (see above)

(5) Existence of an Additive Identity –there is a single unique vector $\mathbf{0}$ in V that when added to any vector \mathbf{v} in V gives back the vector \mathbf{v} :

$$\boxed{\mathbf{v} + \mathbf{0} = \mathbf{v}}$$

This vector is called the zero vector, and for C^N and R^N is the vector containing N zeros. This axiom just says that every vector space must include the zero vector.

(6) Existence of Negative Vector: For every vector \mathbf{v} in V there exists another element, say $\tilde{\mathbf{v}}$, such that

$$\boxed{\tilde{\mathbf{v}} + \mathbf{v} = \mathbf{0}}$$

that is, it is the negative of \mathbf{v} : $\boxed{\tilde{\mathbf{v}} = -1 \cdot \mathbf{v}}$

Subspace: a subset of vectors in the space that itself is closed under vector addition and scalar multiplication (using the same set of scalars) is called a subspace of the original vector space.

- Examples:
- 1. R^2 is a subspace of R^3 .
 - 2. Any line passing through the origin in R^2 is a subspace of R^2 – Verify it!
 - 3. R^2 is NOT a subspace of C^2 because R^2 isn't closed under complex scalars

So, we have a set of vectors plus an “arithmetic structure” on that set of vectors and we call that a vector space. Now we add some more structure to the vector space... Let’s add some “geometric structure” by considering the length of a vector.

Length of a Vector (Vector Norm): For any $\mathbf{v} \in C^N$ we define its length (or norm) to be

$$\|\mathbf{v}\|_2 = \sqrt{\sum_{i=1}^N |v_i|^2}$$

This is called the “Two Norm” or the “Euclidian Norm”; there are other types of norms, but this is the one we’ll need.

Properties of Vector Norm:

- 1. $\|\alpha\mathbf{v}\|_2 = |\alpha| \|\mathbf{v}\|_2$

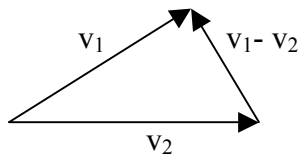
2. $\|\alpha \mathbf{v}_1 + \beta \mathbf{v}_2\|_2 \leq \alpha \|\mathbf{v}_1\|_2 + \beta \|\mathbf{v}_2\|_2$
3. $\|\mathbf{v}\|_2 < \infty \quad \forall \mathbf{v} \in C^N$
4. $\|\mathbf{v}\|_2 = 0$ iff $\mathbf{v} = \mathbf{0}$ (the zero vector has all elements equal to 0)

So, we now have “length” ideas; now we build in distance

Distance Between Vectors: the distance between two vectors in a vector space with the two norm is defined by:

$$d(\mathbf{v}_1, \mathbf{v}_2) = \|\mathbf{v}_1 - \mathbf{v}_2\|_2$$

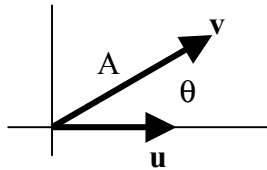
Note that $d(\mathbf{v}_1, \mathbf{v}_2) = 0$ iff $\mathbf{v}_1 = \mathbf{v}_2$



Now, more geometry comes because the “angle” between vectors can be defined in terms of the inner product...

Inner Product Between Vectors:

Motivate the idea in \mathbb{R}^2 :



$$\mathbf{v} = \begin{bmatrix} A \cos \theta \\ A \sin \theta \end{bmatrix} \quad \mathbf{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\sum_{i=1}^2 u_i v_i = 1 \cdot A \cos \theta + 0 \cdot A \sin \theta = A \cos \theta = 0 \quad \text{iff } \theta = 90^\circ$$

A measure of the angle between the vectors

In general C^N : $\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^N u_i v_i^*$

Properties of Inner Products:

1. $\langle \alpha \mathbf{u}, \mathbf{v} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle$
 $\langle \mathbf{u}, \beta \mathbf{v} \rangle = \beta^* \langle \mathbf{u}, \mathbf{v} \rangle$
2. $\langle \mathbf{u}, \mathbf{v} + \mathbf{z} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{z} \rangle$
 $\langle \mathbf{u} + \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle$
3. $\|\mathbf{v}\|_2^2 = \langle \mathbf{v}, \mathbf{v} \rangle$
4. $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|_2 \|\mathbf{v}\|_2$ Schwarz Inequality
5. $\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|_2 \|\mathbf{v}\|_2}$

- (i) lies between -1 and 1;
- (ii) measures directional alikeness of \mathbf{u} and \mathbf{v}
 - = +1 when \mathbf{u} and \mathbf{v} point in the same direction
 - = 0 when \mathbf{u} and \mathbf{v} are a “right angle”
 - = -1 when \mathbf{u} and \mathbf{v} point in opposite directions

6. Two vectors \mathbf{u} and \mathbf{v} for which $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ are called **orthogonal** vectors

If in addition, they each have unit length ($\|\mathbf{u}\|_2 = 1$ $\|\mathbf{v}\|_2 = 1$) they are said to be **orthonormal**

Building Vectors from Other Vectors

Can we find a set of “prototype” vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_M\}$ from which we can build all other vectors in some given vector space?

We’d like to find a set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_M\}$ such that we can write any vector in terms of a linear combination of these vectors:

For example, if we had $\mathbf{v} \in C^N$, we could build it using $\sum_{k=1}^M \alpha_k \mathbf{v}_k$

... and if we had $\mathbf{u} \in C^N$, we could build it using $\sum_{k=1}^M \beta_k \mathbf{v}_k$

The only difference between how we build these two vectors is the coefficients that we choose (α 's for \mathbf{v} and β 's for \mathbf{u}). Thus, the vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_M\}$ act as the ingredients in the recipe and the α 's and β 's act as the amounts of the ingredients we need to build a certain vector.

Thus, what we want to be able to do is get any vector we want by changing the amounts. To do this requires that the set of prototype vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_M\}$ satisfy certain conditions. We'd also like to make sure that the set of prototype vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_M\}$ is as small in number as possible (no sense using more vectors here than is necessary, right?). To answer this question we need the idea of linear independence...

Linear Independence: A set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_M\}$ is said to be linear independent if there is no vector in it that can be written as a linear combination of the others.

If there were a vector that could be written in terms of the others, then this vector is not really needed if we are trying to find a minimal set of prototype vectors because any contribution it might make to build a desired vector could be obtained using the vectors that can make it! For example, say that we have a set of four vectors $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ and lets say that we know that we can build \mathbf{v}_2 from \mathbf{v}_1 and \mathbf{v}_3 according to $\mathbf{v}_2 = 2\mathbf{v}_1 + 3\mathbf{v}_3$. Now lets say we are interested in finding how to build some vector \mathbf{u} from the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ according to

$$\mathbf{u} = \sum_{k=1}^4 \alpha_k \mathbf{v}_k = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 + \alpha_4 \mathbf{v}_4$$

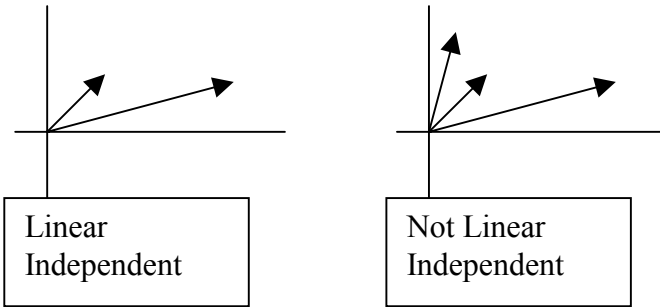
But we know that \mathbf{v}_2 can be written in terms of \mathbf{v}_1 and \mathbf{v}_3 so that we have:

$$\begin{aligned} \mathbf{u} &= \alpha_1 \mathbf{v}_1 + \alpha_2 [2\mathbf{v}_1 + 3\mathbf{v}_3] + \alpha_3 \mathbf{v}_3 + \alpha_4 \mathbf{v}_4 \\ &= (\alpha_1 + 2\alpha_2) \mathbf{v}_1 + (\alpha_3 + 3\alpha_2) \mathbf{v}_3 + \alpha_4 \mathbf{v}_4 \\ &= \hat{\alpha}_1 \mathbf{v}_1 + \hat{\alpha}_3 \mathbf{v}_3 + \alpha_4 \mathbf{v}_4 \end{aligned}$$

Thus, we see that \mathbf{v}_2 is not needed at all.

Linear independence is a property that allows us to check to see if we have too many vectors in a set of vectors we are considering as a possible prototype set – in other words, it allows us to check to see if the proposed set has any redundancy in it.

Examples of Linear Independence:



It is clear that:

1. In C^N or R^N we can have no more than N linear independent vectors.
2. Any set of mutually orthogonal vectors is linear independent (a set of vectors is mutually orthogonal if all pairs are orthogonal).

Span of a Set of Vectors: A set of vectors $\{v_1, \dots, v_M\}$ is said to span a vector space V if it is possible to write any vector v in V using only a linear combination of vectors from the set $\{v_1, \dots, v_M\}$:

$$v = \sum_{k=1}^M \alpha_k v_k$$

This is a check to see if there are enough vectors in the proposed prototype set to build all possible vectors.

Thus, because what we want is a smallest such prototype set, we see that we need a set that spans the space and is linearly independent. This leads to the following definition.

Basis of a Vector Space: A basis of a vector space is a set of linear independent vectors that span the space.

Orthonormal (ON) Basis: If a basis of a vector space contains vectors that are orthonormal to each other (all pairs of basis vectors are orthogonal and each basis vector has unit length).

Fact: Any set of N linearly independent vectors in C^N (R^N) is a basis of C^N (R^N).

Dimension of a Vector Space: The number of vectors in any basis for a vector space is said to be the dimension of the space. Thus, C^N and R^N each have dimension of N .

Fact: For a given basis $\{v_1, \dots, v_N\}$, the expansion of a vector v in V is unique. That is, for each v there is only one unique set of coefficients $\{\alpha_1, \dots, \alpha_N\}$ such that

$$\mathbf{v} = \sum_{k=1}^N \alpha_k \mathbf{v}_k$$

In other words, this expansion or decomposition is unique. Thus, for a given basis we can make a one-to-one correspondence between the vector \mathbf{v} and the expansion coefficients $\{\alpha_1, \dots, \alpha_N\}$. Note we can write the expansion coefficients as a vector, too:

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix}$$

Coefficients as Transformed Values: Thus, we can think of the expansion as being a mapping (or transformation) from vector \mathbf{v} to vector $\boldsymbol{\alpha}$. Knowing the coefficient vector $\boldsymbol{\alpha}$ provides complete knowledge of vector \mathbf{v} as long as we know the basis set to be used with $\boldsymbol{\alpha}$. We can think of the process of computing the coefficients as applying a transform to the original vector. We can view this transform as taking us from the original vector space into a new vector space made from the coefficient vectors of all the vectors in the original space.

Fact: For any given vector space there are an infinite number of possible basis sets. The coefficients with respect to any of them provides complete information about a vector; however, some of them provide more insight into the vector and are therefore more useful for certain signal processing tasks than others.

Example for \mathbf{R}^3

We know that the dimension of this space is $N=3$ and that we therefore need three vectors in any basis for \mathbf{R}^3 . We could pick the simplest basis:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (\text{Verify that this is indeed linearly independent})$$

Note also that this particular basis is an ON basis because each pairs of vectors is orthogonal and each vector has unit length.

Now, we can write any \mathbf{R}^3 vector $\mathbf{v} = [\alpha \ \beta \ \gamma]^T$ by a linear combination of the basis vectors \mathbf{e}_1 , \mathbf{e}_2 , and \mathbf{e}_3 and it is easy to verify that the coefficients for this basis are nothing more than the vector's elements themselves:

$$\mathbf{v} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Thus, this basis really doesn't give any more insight than the vector itself. To explore this farther, let's say that we have a scenario where we know that any vector we would have will follow the specific form $\mathbf{v} = [\alpha \quad 2\alpha \quad \gamma]^T$. Then perhaps a better basis to choose would be

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \quad (\text{Verify that this is an orthogonal basis – it can be normalized to give an ON basis.})$$

Then using this basis we can transform any vector of the form $\mathbf{v} = [\alpha \quad 2\alpha \quad \gamma]^T$ into the corresponding set coefficients $\{\alpha, \gamma, 0\}$; thus, for this limited class of vectors we need only 2 numbers to describe them rather than three as would be needed for the basis set of $\mathbf{e}_1, \mathbf{e}_2,$ and \mathbf{e}_3 . For this case we have found a basis that exploits structure in the set of vectors expected and the result is a more efficient means of representing the vector. Note that this works here because the expected form of the vectors limits them to a 2 dimensional subspace of \mathbb{R}^3 that is spanned by the first two vectors in the “v” basis. In signal processing we try to do the same thing: find a basis that yields an efficient basis for the types of vectors that we are trying to deal with.

DFT from Basis Viewpoint:

Consider that we have a discrete-time signal $x[n]$ for $n = 0, 1, \dots, N-1$. We know that we can compute the DFT of this using

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi kn / N}$$

where $X[k]$ for $k = 0, 1, \dots, N-1$ are the DFT coefficients. We also know that we can build (i.e., get back) the original signal using the inverse DFT (IDFT) according to

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{j2\pi kn / N}$$

Now let's view this from the vector space perspective we've been developing here. Let the vector \mathbf{x} consist of the original signal samples:

$$\mathbf{x} = [x[0] \quad x[1] \quad \dots \quad x[N-1]]^T$$

and define a set of basis vectors from the complex exponentials that are used in the IDFT sum – the k^{th} one looks like this:

$$\mathbf{d}_k = \begin{bmatrix} e^{j2\pi k 0 / N} \\ e^{j2\pi k 1 / N} \\ \vdots \\ e^{j2\pi k (N-1) / N} \end{bmatrix}$$

where the values down the column follow the running n variable; thus the k^{th} vector viewed as a signal is N samples of a complex sinusoid of frequency $2\pi k / N$. Thus, the set of N vectors \mathbf{d}_k for $k = 0$ to $N-1$ are

$$\mathbf{d}_0 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \mathbf{d}_1 = \begin{bmatrix} 1 \\ e^{j2\pi 1 / N} \\ \vdots \\ e^{j2\pi 1 (N-1) / N} \end{bmatrix} \quad \mathbf{d}_2 = \begin{bmatrix} 1 \\ e^{j2\pi 2 / N} \\ \vdots \\ e^{j2\pi 2 (N-1) / N} \end{bmatrix} \quad \dots \quad \mathbf{d}_{N-1} = \begin{bmatrix} 1 \\ e^{j2\pi (N-1) / N} \\ \vdots \\ e^{j2\pi (N-1)(N-1) / N} \end{bmatrix}$$

Using the standard results that are used for DFT/IDFT analysis it is easy to show that these vectors are indeed orthogonal and thus, because there are N of them, form a basis for \mathbb{C}^N . Thus, we can view the IDFT equation above as an expansion of the signal vector \mathbf{x} in terms of this complex sinusoid basis:

$$\mathbf{x} = \frac{1}{N} \sum_{k=0}^{N-1} X[k] \mathbf{d}_k,$$

from which it can be seen that the basis coefficients are proportional to the DFT coefficients. Thus, if a signal is a sinusoid (or contains significant sinusoidal components) then this basis provides an efficient representation. Note that our view of using the coefficient vector as a complete characterization of the vector is consistent with the usual view of using the DFT coefficients as a complete characterization of the signal.

Usefulness of ON Basis

Note: any orthogonal basis can be made into an ON basis by scaling each vector in the basis to have unit norm.

What's So Good About an ON Basis?: Let's See! Given any basis $\{\mathbf{v}_1, \dots, \mathbf{v}_N\}$ we can write any \mathbf{v} in V as

$$\mathbf{v} = \sum_{k=1}^N \alpha_k \mathbf{v}_k$$

with unique coefficients. Now, we ask: Given the vector \mathbf{v} how we find the α 's?
 In general – hard!
 For ON basis – easy!! Here is how.

If $\{\mathbf{v}_1, \dots, \mathbf{v}_N\}$ is an ON basis, then

$$\begin{aligned} \langle \mathbf{v}, \mathbf{v}_i \rangle &= \left\langle \left[\sum_{j=1}^N \alpha_j \mathbf{v}_j \right], \mathbf{v}_i \right\rangle \\ &= \sum_{j=1}^N \alpha_j \underbrace{\langle \mathbf{v}_j, \mathbf{v}_i \rangle}_{\delta_{[j-i]}} \\ &= \alpha_i \end{aligned}$$

Thus, the i^{th} coefficient is found by taking the inner product between the vector \mathbf{v} and the i^{th} basis vector:

$$\boxed{\alpha_i = \langle \mathbf{v}, \mathbf{v}_i \rangle} \quad \text{EASY!!!!}$$

The other thing that is good about an ON basis is that they preserve inner products and norms. Let \mathbf{u} and \mathbf{v} be two vectors that are expanded with respect to some ON basis to give the coefficient vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, respectively. Then,

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle$ (Preserves Inner Products)
2. $\|\mathbf{u}\|_2 = \|\boldsymbol{\alpha}\|_2$ and $\|\mathbf{v}\|_2 = \|\boldsymbol{\beta}\|_2$ (Preserves Norms)

Actually, the second of these is a special case of the first. These are known as Parseval's theorem. The first allows us to characterize the closeness of two vectors in terms of the closeness of their respective coefficient vectors. The second allows us to measure the size of a vector in terms of the size of its coefficient vector.

So... What's so good about an ON basis? It provides an easy way to compute the coefficients and it ensures a link between the coefficient space "geometry" and the original space "geometry".

DFT Coefficients as Inner Product Results:

Now lets see how these ideas relate to the DFT. We've already seen that we can interpret the N-pt. IDFT as an expansion of the signal vector in terms of the linear combination of N ON vectors \mathbf{d}_k defined above. (Note that for the IDFT case this is just saying that we are building the signal out of a sum of complex sinusoids.) From the above theory, we now know how to compute the required coefficients for any ON expansion so we should see what this idea gives for the IDFT case. Our main general result above was that the expansion coefficients are found from the inner products of the

vector to be expanded and the various ON basis vectors: so for the IDFT case the coefficients are $\langle \mathbf{x}, \mathbf{d}_k \rangle$. But, we know from standard DFT theory that the coefficients of the IDFT expansion are just the DFT values $X[k]$. Combining these two points of view gives

$$\boxed{X[k] = \langle \mathbf{x}, \mathbf{d}_k \rangle}$$

Let's take a look at this and verify that the right-hand side of this is consistent with what we know from standard DFT theory. From vector inner product theory we know that the right side of this is

$$\begin{aligned} \langle \mathbf{x}, \mathbf{d}_k \rangle &= \sum_{n=0}^{N-1} x[n] d_k^*[n] \\ &= \sum_{n=0}^{N-1} x[n] e^{-j2\pi kn/N} \end{aligned}$$

which is the DFT of the signal, so we see that things work as stated. Actually, we have been a little bit loose here – the \mathbf{d}_k vectors as given above are orthogonal but not orthonormal so there is a $1/N$ term in the IDFT equation used back in the “DFT from Basis Viewpoint” section. To really do this using ON basis vectors we would have to put a $1/\sqrt{N}$ term in front of the \mathbf{d}_k vectors, but that would lead to forms that aren't exactly the same as the conventional DFT/IDFT expressions; try as an exercise redoing this development with true ON basis vectors.

Matrices

Matrix: Is an array of numbers organized in rows and columns; here is a 4x4 example

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

It will be helpful sometimes to view a matrix as being built out of its columns; thus, the 4x4 example above could be written as:

$$\mathbf{A} = [\mathbf{a}_1 \mid \mathbf{a}_2 \mid \mathbf{a}_3 \mid \mathbf{a}_4]$$

where $\mathbf{a}_k = \begin{bmatrix} a_{1k} \\ a_{2k} \\ a_{3k} \\ a_{4k} \end{bmatrix}$

Matrix as Transform: Our main view of matrices will be as things that transform or map a vector from one space to another. Consider the 4x4 example matrix above. We could use that to transform one 4 dimensional vector space V into another one U as follows. Let \mathbf{v} be in V and compute the product $\mathbf{A}\mathbf{v}$ and call it vector \mathbf{u} :

$$\begin{aligned} \mathbf{u} = \mathbf{A}\mathbf{v} &= [\mathbf{a}_1 \mid \mathbf{a}_2 \mid \mathbf{a}_3 \mid \mathbf{a}_4]\mathbf{v} \\ &= v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3 + v_4\mathbf{a}_4 \end{aligned}$$

as computed according to the definition of matrix-vector multiplication. Now what vector space does \mathbf{u} lie in? It clearly is built from the columns of matrix \mathbf{A} ; therefore, it can only be a vector that lies in the span of the set of vectors that make up the columns of \mathbf{A} . Exactly which vector in this space \mathbf{v} gets mapped to clearly depends on what the values of the elements of \mathbf{v} are, since they act as the “building” coefficients for the columns of \mathbf{A} when \mathbf{u} is built.

If we apply \mathbf{A} to all the vectors in V we get a collection of vectors that are in a new space called U .

Fact: If the mapping matrix \mathbf{A} is square and its columns are linearly independent then
 (i) the space that vectors in V get mapped to (i.e., U) has the same dimension as V
 (ii) this mapping is reversible (i.e., invertible); there is an inverse matrix \mathbf{A}^{-1} such that $\mathbf{v} = \mathbf{A}^{-1}\mathbf{u}$.

Matrix View of Basis

Recall: For a general basis $\{\mathbf{v}_1, \dots, \mathbf{v}_N\}$ we can expand a vector in the space according to

$$\mathbf{v} = \sum_{k=1}^N \alpha_k \mathbf{v}_k$$

Another view: Consider the $N \times N$ matrix \mathbf{V} whose columns are the basis vectors:

$$\mathbf{V} = [\mathbf{v}_1 \mid \dots \mid \mathbf{v}_N]$$

And consider the $N \times 1$ vector of coefficients:

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix}$$

Then we can rewrite the expansion using matrix multiplication:

$$\mathbf{v} = \mathbf{V}\boldsymbol{\alpha} = \underbrace{[\mathbf{v}_1 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_N]}_{N \times N \text{ matrix}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix}$$

We can now take three views of this:

View #1: Vector \mathbf{v} is a linear combination of the columns of basis matrix \mathbf{V} .

View #2: Matrix \mathbf{V} maps vector $\boldsymbol{\alpha}$ into vector \mathbf{v}

View #3: Is there a matrix, say $\boldsymbol{\Gamma}$, that maps vector \mathbf{v} into vector $\boldsymbol{\alpha}$?

Aside: If a matrix \mathbf{A} is square and has linearly independent columns, then \mathbf{A}^{-1} exists such that $\mathbf{A} \mathbf{A}^{-1} = \mathbf{A}^{-1} \mathbf{A} = \mathbf{I}$ where \mathbf{I} is the identity matrix having 1's on the diagonal and zeros elsewhere.

Well... because \mathbf{V} is a basis matrix, it is square and its columns are linearly independent. Thus, its inverse exists and we have:

$$\boldsymbol{\alpha} = \mathbf{V}^{-1} \mathbf{v}$$

So we now have a way to go back and forth between the vector \mathbf{v} and its coefficient vector $\boldsymbol{\alpha}$.

Basis Matrix for ON Basis: When the basis used to form the matrix is ON then we get a special structure that arises (this is connected to the inner product result we derived for computing the coefficients of an ON basis).

Result: The inverse of the ON basis matrix \mathbf{V} is \mathbf{V}^H , where the superscript H denotes hermitian transpose, which consists of transposing the matrix and conjugating the elements. Let's see where this comes from. We need to show that $\mathbf{V} \mathbf{V}^H = \mathbf{I}$. Consider the form for $\mathbf{V} \mathbf{V}^H$:

$$\mathbf{V}\mathbf{V}^H = \begin{bmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \langle \mathbf{v}_1, \mathbf{v}_2 \rangle & \cdots & \langle \mathbf{v}_1, \mathbf{v}_N \rangle \\ \langle \mathbf{v}_2, \mathbf{v}_1 \rangle & \langle \mathbf{v}_2, \mathbf{v}_2 \rangle & \cdots & \langle \mathbf{v}_2, \mathbf{v}_N \rangle \\ \vdots & & \ddots & \vdots \\ \langle \mathbf{v}_N, \mathbf{v}_1 \rangle & \langle \mathbf{v}_N, \mathbf{v}_2 \rangle & \cdots & \langle \mathbf{v}_N, \mathbf{v}_N \rangle \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \mathbf{I}$$

where we have used the fact that the vectors are ON.

Thus, for the ON basis case, the inverse basis matrix is easy to find from the basis matrix:

$$\boxed{\mathbf{V}^{-1} = \mathbf{V}^H}$$

Such a matrix (i.e., one whose inverse is just its hermitian transpose) is called a unitary matrix; in other words, a unitary matrix is one that satisfies $\mathbf{V}\mathbf{V}^H = \mathbf{V}^H\mathbf{V} = \mathbf{I}$. As a note, for the real-valued matrix case the idea of unitary matrix becomes what is called an orthogonal matrix for which $\mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}$ and $\mathbf{V}^{-1} = \mathbf{V}^T$; that is, for real-valued matrices the hermitian transpose is the same as the regular transpose.

Two Properties of Unitary Matrices:

1. They preserve norms: Let \mathbf{U} be a unitary matrix and let $\mathbf{y} = \mathbf{U}\mathbf{x}$.

$$\text{Then } \|\mathbf{y}\| = \|\mathbf{x}\|.$$

2. They preserve inner products: Let \mathbf{U} be a unitary matrix and let $\mathbf{y}_1 = \mathbf{U}\mathbf{x}_1$ and $\mathbf{y}_2 = \mathbf{U}\mathbf{x}_2$.

$$\text{Then } \langle \mathbf{y}_1, \mathbf{y}_2 \rangle = \langle \mathbf{x}_1, \mathbf{x}_2 \rangle$$

That is norms and inner products are preserved by the unitary matrix as it transforms into the new space.

This is the same thing as the preservation properties of ON basis.

DFT from Unitary Matrix Viewpoint:

Consider that we have a discrete-time signal $x[n]$ for $n = 0, 1, \dots, N-1$. We know that we can compute the DFT of this using

$$X[k] = \sum_{n=0}^{N-1} x[n]e^{-j2\pi kn/N}$$

where $X[k]$ for $k = 0, 1, \dots, N-1$ are the DFT coefficients. We also know that we can build (i.e., get back) the original signal using the inverse DFT (IDFT) according to

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{j2\pi kn/N}$$

Now lets view this from the unitary matrix perspective we've been developing here. Let the vector \mathbf{x} consist of the original signal samples:

$$\mathbf{x} = [x[0] \quad x[1] \quad \dots \quad x[N-1]]^T$$

and define a set of basis vectors from the complex exponentials that are used in the IDFT sum:

$$\mathbf{d}_k = \begin{bmatrix} e^{j2\pi k 0/N} \\ e^{j2\pi k 1/N} \\ \vdots \\ e^{j2\pi k (N-1)/N} \end{bmatrix}$$

where the values down the column follow the running n variable; thus the k^{th} vector is a complex sinusoid of frequency $2\pi k / N$. Thus, the set of N vectors \mathbf{d}_k for $k = 0$ to $N-1$ are

$$\mathbf{d}_0 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \mathbf{d}_1 = \begin{bmatrix} 1 \\ e^{j2\pi 1 \cdot 1/N} \\ \vdots \\ e^{j2\pi 1(N-1)/N} \end{bmatrix} \quad \mathbf{d}_2 = \begin{bmatrix} 1 \\ e^{j2\pi 2 \cdot 1/N} \\ \vdots \\ e^{j2\pi 2(N-1)/N} \end{bmatrix} \quad \dots \quad \mathbf{d}_{N-1} = \begin{bmatrix} 1 \\ e^{j2\pi(N-1) \cdot 1/N} \\ \vdots \\ e^{j2\pi(N-1)(N-1)/N} \end{bmatrix}$$

Using the standard results that are used for DFT/IDFT analysis it is easy to show that these vectors are indeed orthogonal and thus, because there are N of them, form a basis for \mathbb{C}^N ; they aren't orthonormal, as mentioned above, so we have to account for that in our usage here. We'll see that we can view the IDFT equation above as *nearly* unitary matrix – where the “*nearly*” comes from the fact that vectors aren't normalized.

Recall that we've seen that we can write the IDFT as follows

$$\mathbf{x} = \frac{1}{N} \sum_{k=0}^{N-1} X[k] \mathbf{d}_k .$$

From this motivation we define the matrix \mathbf{D} as follows:

$$\mathbf{D} = [\mathbf{d}_0 \mid \mathbf{d}_1 \mid \dots \mid \mathbf{d}_{N-1}] = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & e^{j2\pi 1/N} & e^{j2\pi 2/N} & \dots & e^{j2\pi(N-1)/N} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & e^{j2\pi 1(N-1)/N} & e^{j2\pi 2(N-1)/N} & \dots & e^{j2\pi(N-1)(N-1)/N} \end{bmatrix}$$

from which we can write the IDFT as

$$\mathbf{x} = \frac{1}{N} \mathbf{D} \tilde{\mathbf{x}}$$

where we have used the vector $\tilde{\mathbf{x}}$ to denote the vector of DFT coefficients $X[k]$. This is the “Matrix View of a Basis” applied to the IDFT.

Now to get the DFT coefficients from the signal vector, we use the fact that the matrix \mathbf{D} is (nearly) unitary so that we have that

$$\frac{1}{N} \mathbf{D}^H \mathbf{D} = \mathbf{I}$$

so that we get

$$\begin{aligned} \mathbf{D}^H \mathbf{x} &= \frac{1}{N} \mathbf{D}^H \mathbf{D} \tilde{\mathbf{x}} \\ &= \mathbf{I} \tilde{\mathbf{x}} = \tilde{\mathbf{x}} \end{aligned}$$

or in other words, to compute the vector of DFT coefficients $\tilde{\mathbf{x}}$ all we have to do is

$$\tilde{\mathbf{x}} = \mathbf{D}^H \mathbf{x}$$

You should be able to verify that this matrix operation does indeed give the DFT coefficients.

Matrix as Transform

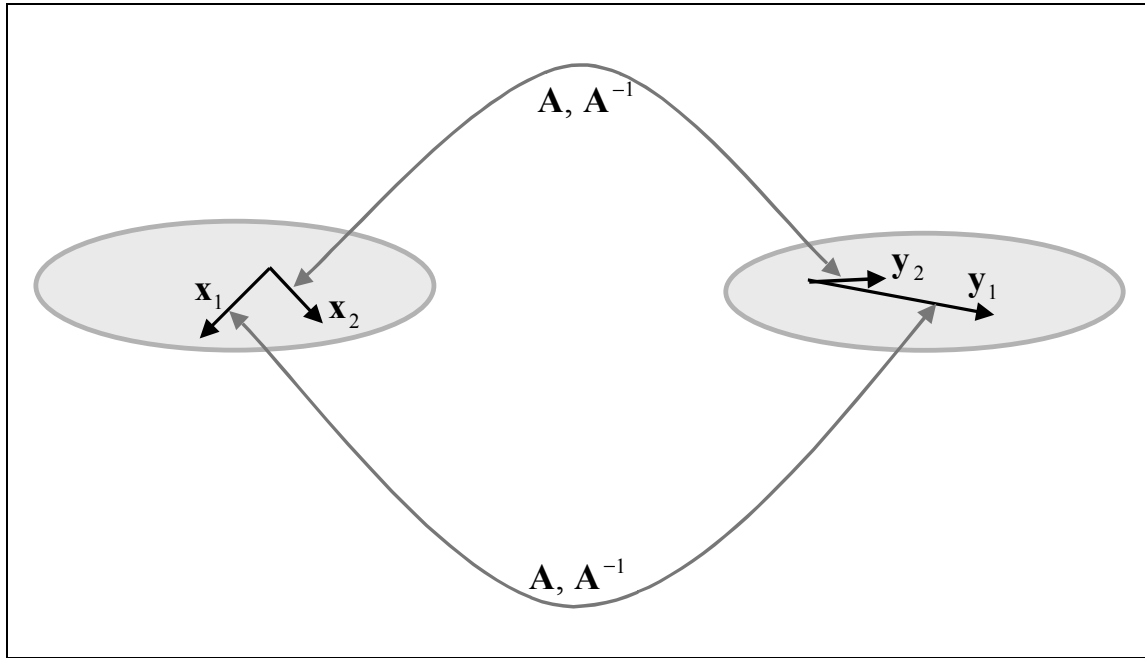
What we just saw for the DFT was that we can view the Discrete Fourier **transform** as a matrix multiplying a vector. This is a VERY useful viewpoint for all sorts of signal transforms. In general we can view any linear transform in terms of some matrix (say \mathbf{A}) operating on a signal vector (say \mathbf{x}) by pre-multiplication to give another vector (say \mathbf{y}):

$$\mathbf{y} = \mathbf{A} \mathbf{x}$$

Given what we seen above we can view this as mapping the vector \mathbf{x} into a new vector space, where that new space is the span of the columns of matrix \mathbf{A} . If \mathbf{A} is square and invertible, then we can get \mathbf{x} back from \mathbf{y} by operating on \mathbf{y} with the inverse of \mathbf{A} :

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$$

Thus, we see that we can think of \mathbf{A} and its inverse as mapping back and forth between two vector spaces as shown below:



This figure shows two vectors \mathbf{x}_1 and \mathbf{x}_2 getting mapped into \mathbf{y}_1 and \mathbf{y}_2 . The double direction arrows labeled with \mathbf{A} and its inverse show that it is possible to go back forth between these two spaces (or domains). Note that we have seen that unitary matrices do this mapping in such a way that the sizes of vectors and the orientation between vectors is not changed – clearly this picture shows a non-unitary matrix because the sizes of the \mathbf{y} 's are not the same as the sizes of the \mathbf{x} 's nor are the \mathbf{y} 's oriented the same as the \mathbf{x} 's.

While this viewpoint is very useful; it is important to always keep in your mind that there is an alternative, basis viewpoint – namely, that the transform results \mathbf{y} gives the coefficients that are used to write the recipe for how to build the \mathbf{x} vector as a linear combination of the columns of \mathbf{A}^{-1} . Often in signal processing we choose the matrix \mathbf{A} to ensure that the computed \mathbf{y} vector gives us useful insight to the vector \mathbf{x} that was not available directly from \mathbf{x} – this is what is done in the DFT and other transforms used for data compression, etc. For data compression, if we choose \mathbf{A} such the resulting vector \mathbf{y} has mostly small (or even better, zero) elements, then we can use that to advantage to reduce the amount of data that we need to send. For DFT, we often use it to analyze sinusoidal signals – that's because a sinusoid gives a large value for the element of \mathbf{y} in the position corresponding to the basis vector frequency that is closest to the signal's frequency; the other elements of \mathbf{y} are likely to be small, thus we can use this to detect the presence of a sinusoid and measure its frequency.