Ch. 12 Review of Matrices and Vectors

Vectors & Vector Spaces

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

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix} = [v_1 \cdots v_N]^T$$

Transpose

Definition of Vector Addition: Add element-by-element

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

Definition of Scalar: A real or complex number.

If the vectors of interest are complex valued then the set of scalars is taken to be complex numbers; if the vectors of interest are real valued then the set of scalars is taken to be real numbers.

Multiplying a Vector by a Scalar :

$$\mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix} \qquad\qquad \alpha\mathbf{a} = \begin{bmatrix} \alpha a_1 \\ \vdots \\ \alpha a_N \end{bmatrix}$$

…changes the vector's length if $|\alpha| \neq 1$

… "reverses" its direction if $\alpha < 0$

Arithmetic Properties of Vectors: vector addition and scalar multiplication exhibit the following properties pretty much like the real numbers do

Let **x**, **y**, and **z** be vectors of the same dimension and let α and β be scalars; then the following properties hold:

1. Commutativity

$$\boxed{\begin{array}{l} \mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x} \\ \\ \alpha \mathbf{x} = \mathbf{x}\alpha \end{array}}$$

2. Associativity

$$\boxed{\begin{array}{l} (\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{y} + (\mathbf{x} + \mathbf{z}) \\ \\ \alpha(\beta \mathbf{x}) = (\alpha\beta)\mathbf{x} \end{array}}$$

3. Distributivity

$$\boxed{\begin{array}{l} \alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y} \\ \\ (\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x} \end{array}}$$

4. Scalar Unity & Scalar Zero

$$\boxed{\begin{array}{l} 1\mathbf{x} = \mathbf{x} \\ \\ 0\mathbf{x} = \mathbf{0}, \quad \text{where } \mathbf{0} \text{ is the zero vector of all zeros} \end{array}}$$

<u>Definition of a Vector Space</u>: A set V of N-dimensional vectors (with a corresponding set of scalars) such that the set of vectors is:

(i) "closed" under vector addition
(ii) "closed" under scalar multiplication

In other words:

- addition of vectors – gives another vector <u>in the set</u>
- multiplying a vector by a scalar – gives another vector <u>in the set</u>

Note: this means that ***<u>ANY</u>*** "<u>linear combination</u>" of vectors in the space results in a vector in the space…

If $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are all vectors in a given vector space V, then

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 = \sum_{i=1}^{3} \alpha_i \mathbf{v}_i$$

is also in the vector space V.

<u>Axioms of Vector Space</u>: If V is a set of vectors satisfying the above definition of a vector space then it satisfies the following axioms:

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** – any vector space V must have a zero vector
6. **Existence of Negative Vector**: For every vector $\mathbf{v}$ in V its negative must also be in V

So… a vector space is nothing more than a set of vectors with an "<u>arithmetic</u> structure"

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

Examples:

1. The space R^2 is a subspace of R^3.
2. Any plane in R^3 that passes through the origin is a subspace
3. Any line passing through the origin in R^2 is a subspace of R^2
4. The set R^2 is NOT a subspace of C^2 because R^2 isn't closed under complex scalars (a subspace must retain the original space's set of scalars)

Geometric Structure of Vector Space

Length of a Vector (Vector Norm): For any vector **v** in C^N we define its length (or "norm") to be

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

Properties of Vector Norm:

$$\boxed{\|\alpha\mathbf{v}\|_2 = |\alpha| \, \|\mathbf{v}\|_2}$$

$$\boxed{\|\alpha\,\mathbf{v}_1 + \beta\,\mathbf{v}_2\|_2 \le |\alpha|\|\mathbf{v}_1\|_2 + |\beta|\|\mathbf{v}_2\|_2}$$

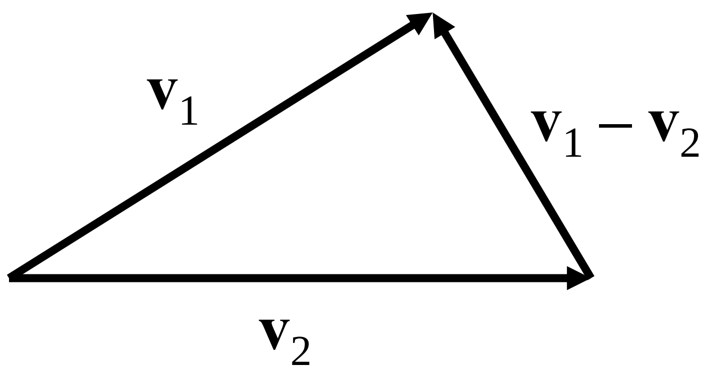
$$\boxed{\|\mathbf{v}\|_2 < \infty \quad \forall \mathbf{v} \in C^N}$$

$$\boxed{\|\mathbf{v}\|_2 = 0 \quad \text{iff} \quad \mathbf{v} = \mathbf{0}}$$

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

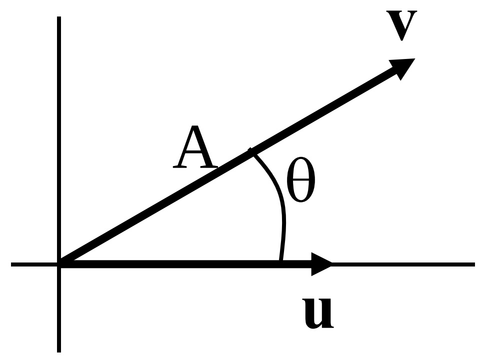
$$\boxed{d(\mathbf{v}_1, \mathbf{v}_2) = \left\| \mathbf{v}_1 - \mathbf{v}_2 \right\|_2}$$

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



Angle Between Vectors & Inner Product:

Motivate the idea in R^2:



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

Note that: $$\sum_{i=1}^{2} u_i v_i = 1\cdot A\cos\theta + 0\cdot A\sin\theta = A\cos\theta$$

Clearly we see that… This gives a measure of the angle between the vectors.

Now we generalize this idea!

Inner Product Between Vectors :

Define the **inner product** between two complex vectors in C^N by:

$$\boxed{< \mathbf{u}, \mathbf{v} >= \sum_{i=1}^{N} u_i v_i^*}$$

Properties of Inner Products:

1. Impact of Scalar Multiplication:

$$< \alpha\mathbf{u}, \mathbf{v} >= \alpha < \mathbf{u}, \mathbf{v} >$$

$$< \mathbf{u}, \beta\mathbf{v} >= \beta^* < \mathbf{u}, \mathbf{v} >$$

2. Impact of Vector Addition:

$$< \mathbf{u}, \mathbf{v} + \mathbf{z} >=< \mathbf{u}, \mathbf{v} > + < \mathbf{u}, \mathbf{z} >$$

$$< \mathbf{u} + \mathbf{w}, \mathbf{v} >=< \mathbf{u}, \mathbf{v} > + < \mathbf{w}, \mathbf{v} >$$

3. Linking Inner Product to Norm:

$$\boxed{\|\mathbf{v}\|_2^2 =< \mathbf{v}, \mathbf{v} >}$$

4. Schwarz Inequality:

$$\boxed{|< \mathbf{u}, \mathbf{v} >| \le \|\mathbf{u}\|_2 \ \|\mathbf{v}\|_2}$$

5. Inner Product and Angle:
(Look back on previous page!)

$$\boxed{\frac{< \mathbf{u}, \mathbf{v} >}{\|\mathbf{u}\|_2 \ \|\mathbf{v}\|_2} = \cos(\theta)}$$

Inner Product, Angle, and Orthogonality :

$$\boxed{\frac{<\mathbf{u},\mathbf{v}>}{\|\mathbf{u}\|_2\,\|\mathbf{v}\|_2} = \cos(\theta)}$$

(i) This lies between –1 and 1;

(ii) It measures <u>directional</u> alikeness of **u** and **v**

= +1 when **u** and **v** point in the same direction

= 0 when **u** and **v** are a "right angle"

= –1 when **u** and **v** point in opposite directions

Two vectors **u** and **v** are said to be **<u>orthogonal</u>** when <**u**,**v**> = 0

If in addition, they <u>each</u> have <u>unit length</u> they are **<u>orthonormal</u>**

Building Vectors From Other Vectors

Can we find a set of "prototype" vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M\}$ from which we can build <u>all</u> other vectors in some given vector space V by using <u>linear combinations</u> of the $\mathbf{v}_i$?

$$\mathbf{v} = \sum_{k=1}^{M} \alpha_k \mathbf{v}_k \qquad\qquad \mathbf{u} = \sum_{k=1}^{M} \beta_k \mathbf{v}_k$$

Same "Ingredients"… just different amounts of them!!!

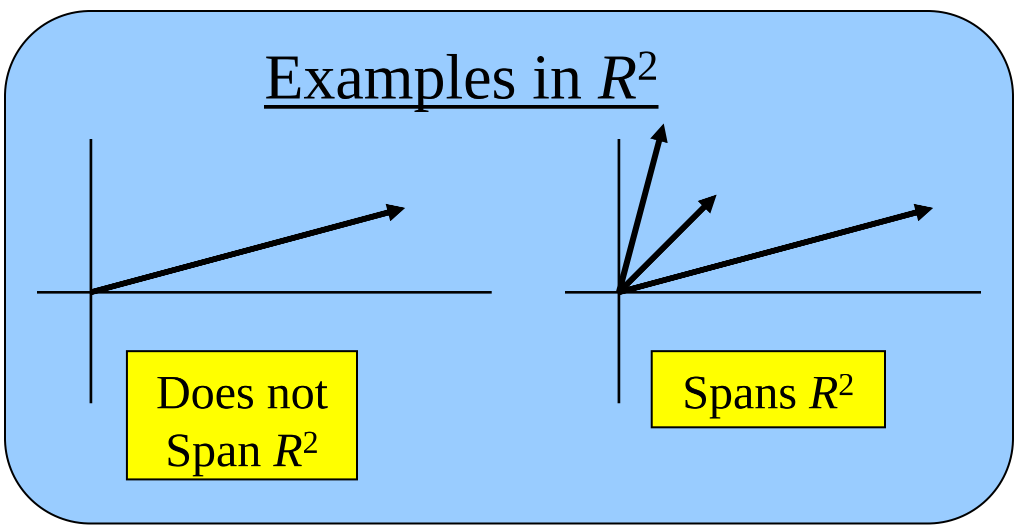
We want to be able to do is get <u>any</u> vector just by changing the amounts… To do this requires that the set of "prototype" vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M\}$ satisfy certain conditions.

We'd also like to have the smallest number of members in the set of "prototype" vectors.

<u>Span of a Set of Vectors</u>: A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M\}$ is said to span the vector space $\boldsymbol{V}$ if it is possible to write each vector $\mathbf{v}$ in $\boldsymbol{V}$ as a linear combination of vectors from the set:

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

This property establishes if there are <u>enough vectors</u> in the proposed prototype set to build all possible vectors in V.



<u>It is clear that</u>:

1. We need at least N vectors to span C^N or R^N but not just any N vectors.

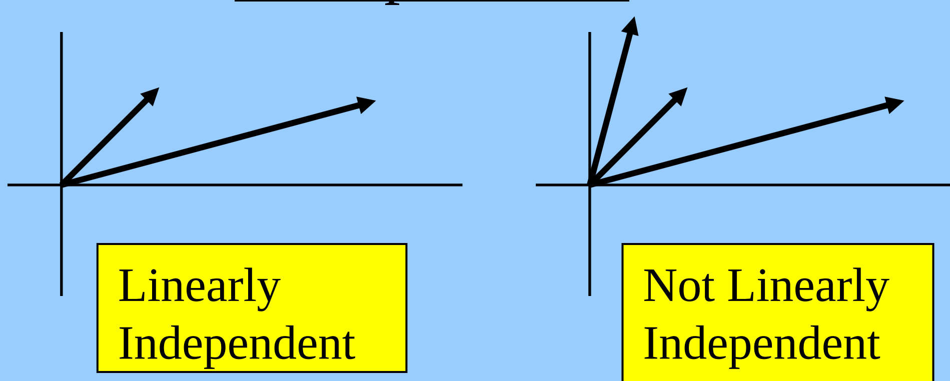
2. Any set of N mutually orthogonal vectors spans C^N or R^N (a set of vectors is mutually orthogonal if all pairs are orthogonal).

Linear Independence: A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M\}$ is said to be linearly <u>in</u>dependent if none of the vectors in it can be written as a linear combination of the others.

If a set of vectors is linearly dependent then there is "redundancy" in the set…it has more vectors than needed to be a "prototype" set!

<u>For example</u>, 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 $\mathbf{v}_3$… then every vector we can build from $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ can also be built from only $\{\mathbf{v}_1, \mathbf{v}_3, \mathbf{v}_4\}$.

<u>Examples in R^2</u>



<u>It is clear that</u>:

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).

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

- "Span" says there are enough vectors to build everything
- "Linear Indep" says that there are not more than needed

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 norm).

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.

Expansion and Transformation

Fact: For a given basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$, the expansion of a vector $\mathbf{v}$ in $\boldsymbol{V}$ is unique. That is, for each $\mathbf{v}$ there is only one, unique set of coefficients $\{\alpha_1, \alpha_2, \ldots, \alpha_N\}$ 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 1-to-1 correspondence between vector $\mathbf{v}$ and the coefficients $\{\alpha_1, \alpha_2, \ldots, \alpha_N\}$.

We can write the coefficients as a vector, too: $\boldsymbol{\alpha} = [\alpha_1 \ \cdots \ \alpha_N]^T$

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix} \quad \xleftrightarrow[\text{1–to–1}]{} \quad \boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix}$$

Expansion can be viewed as a mapping (or transformation) from vector $\mathbf{v}$ to vector $\boldsymbol{\alpha}$.

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 original vectors.

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…

some of them provide more insight into the vector and are therefore more useful for certain signal processing tasks than others.

Often the key to solving a signal processing problem lies in finding the correct basis to use for expanding… this is equivalent to finding the right transform. See discussion coming next linking DFT to these ideas!!!!

<u>DFT from Basis Viewpoint</u>:

If we have a discrete-time signal $x[n]$ for $n = 0, 1, \ldots N$-1

Define vector: $\mathbf{x} = \begin{bmatrix} x[0] & x[1] & \cdots & x[N-1] \end{bmatrix}^T$

Define a orthogonal basis from the exponentials used in the IDFT:

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

Then the IDFT equation can be viewed as an expansion of the signal vector **x** in terms of this complex sinusoid basis:

$$\mathbf{x} = \sum_{k=0}^{N-1} \underbrace{\frac{1}{N} X[k]}_{\alpha_k} \mathbf{d}_k \qquad \boldsymbol{\alpha} = \begin{bmatrix} \frac{X[0]}{N} & \frac{X[1]}{N} & \cdots & \frac{X[N-1]}{N} \end{bmatrix}^T$$

k^{th} coefficient

coefficient vector

Usefulness of an ON Basis

What's So Good About an ON Basis?: Given any basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$ we can write any $\mathbf{v}$ in V as

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

Given the vector $\mathbf{v}$ how do we find the α's?

- In general – hard!   But for ON basis – **easy**!!

If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$ is an ON basis then $\langle \mathbf{v}, \mathbf{v}_i \rangle = \left\langle \left[\sum_{j=1}^{N} \alpha_j \mathbf{v}_j \right], \mathbf{v}_i \right\rangle$

$$= \sum_{j=1}^{N} \alpha_j \underbrace{\langle \mathbf{v}_j, \mathbf{v}_i \rangle}_{\delta[j-i]}$$

$$= \alpha_i$$

$$\boxed{\alpha_i = \langle \mathbf{v}, \mathbf{v}_i \rangle}$$

i^{th} coefficient = inner product with i^{th} ON basis vector

<u>Another Good Thing About an ON Basis</u>: They preserve inner products and norms… (called "isometric"):

If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$ is an ON basis and $\mathbf{u}$ and $\mathbf{v}$ are vectors expanded as

Then….

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

1. $< \mathbf{v}, \mathbf{u} > = < \alpha, \beta >$ (Preserves Inner Prod.)

2. $\|\mathbf{v}\|_2 = \|\alpha\|_2$ and $\|\mathbf{u}\|_2 = \|\beta\|_2$ (Preserves Norms)

So… using an ON basis provides:

- Easy computation via inner products
- Preservation of geometry (closeness, size, orientation, etc.

<u>Example: DFT Coefficients as Inner Products</u>:

Recall: N-pt. IDFT is an expansion of the signal vector in terms of N Orthogonal vectors. Thus

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

$$= \sum_{n=0}^{N-1} x[n] d_k^*[n]$$

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

See "reading notes" for some details about normalization issues in this case

Matrices

Matrix: Is an array of (real or complex) numbers organized in rows and columns.

Here is a 3x4 example:

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

We'll sometimes view a matrix as being built from its columns;

The 3x4 example above could be written as:

$$\mathbf{A} = \left[\mathbf{a}_1 \mid \mathbf{a}_2 \mid \mathbf{a}_3 \mid \mathbf{a}_4\right] \qquad \mathbf{a}_k = \left[a_{1k} \quad a_{2k} \quad a_{3k}\right]^T$$

We'll take two views of a matrix:

1. "Storage" for a bunch of related numbers (e.g., Cov. Matrix)
2. A transform (or mapping, or operator) acting on a vector (e.g., DFT, observation matrix, etc…. as we'll see)

<u>Matrix as Transform</u>: Our main view of matrices will be as "operators" that transform one vector into another vector.

<u>Consider the 3x4 example matrix above</u>. We could use that matrix to transform the 4-dimensional vector **v** into a 3-dimensional vector **u**:

$$\mathbf{u} = \mathbf{A}\mathbf{v} = \left[\mathbf{a}_1 \,|\, \mathbf{a}_2 \,|\, \mathbf{a}_3 \,|\, \mathbf{a}_4\right]\begin{bmatrix} v_1 \\ \\ v_2 \\ \\ v_3 \\ \\ v_4 \end{bmatrix} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3 + v_4\mathbf{a}_4$$

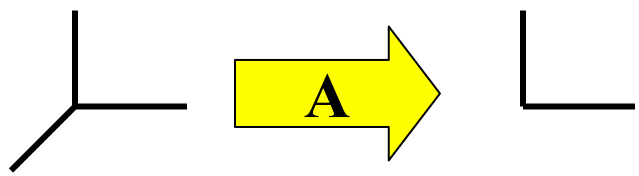
Clearly **u** is built from the columns of matrix **A**; therefore, it must lie in the span of the set of vectors that make up the columns of **A**.

Note that the columns of **A** are 3-dimensional vectors… so is **u.**

<u>Transforming a Vector Space</u>: If we apply **A** to all the vectors in a vector space **V** we get a collection of vectors that are in a new space called **U**.

In the 3x4 example matrix above we transformed a 4-dimensional vector space **V** into a 3-dimensional vector space **U**

A 2x3 real matrix **A** would transform R^3 into R^2 :



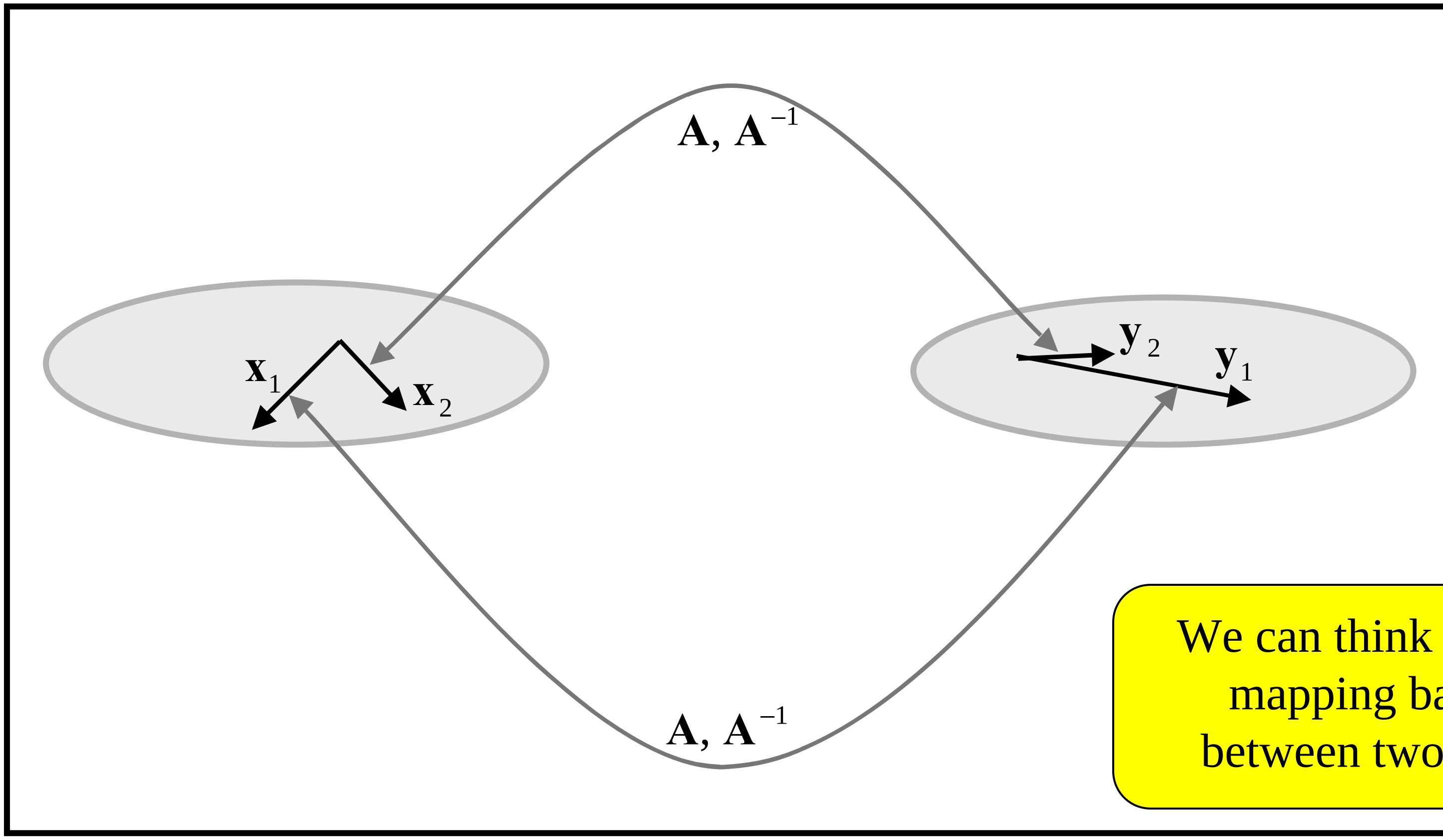
<u>Facts</u>: If the mapping matrix **A** is square <u>and</u> its columns are linearly independent then

(i) the space that vectors in ***V*** get mapped to (i.e., ***U***) has the same dimension as ***V*** *(due to "square" part)*

(ii) this mapping is reversible (i.e., invertible); there is an inverse matrix $\mathbf{A}^{-1}$ such that $\mathbf{v} = \mathbf{A}^{-1}\mathbf{u}$ *(due to "square" & "LI" part)*

Transform = Matrix × Vector: a VERY useful viewpoint for all sorts of signal processing scenarios. In general we can view many linear transforms (e.g., DFT, etc.) in terms of some invertible matrix **A** operating on a signal vector **x** to give another vector **y**:

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



We can think of **A** and **A**⁻¹ as mapping back and forth between two vector spaces

Matrix View & Basis View

Basis Matrix & Coefficient Vector:

Suppose we have a basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$ for a vector space V. Then a vector $\mathbf{v}$ in space V can be written as:

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

Another view of this:

$$\mathbf{v} = \underbrace{\left[\mathbf{v}_1 \,|\, \mathbf{v}_2 \,|\cdots|\, \mathbf{v}_N\right]}_{NxN \text{ matrix}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} \quad \Rightarrow \quad \boxed{\mathbf{v} = \mathbf{V}\boldsymbol{\alpha}}$$

The "Basis Matrix" **V** transforms the coefficient vector into the original vector **v**

<u>Three Views of Basis Matrix & Coefficient Vector</u>:

<u>View #1</u>

Vector **v** is a linear combination of the columns of basis matrix **V**.

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

<u>View #2</u>

Matrix **V** <u>maps</u> vector α into vector **v**.

$$\boxed{\mathbf{v} = \mathbf{V}\boldsymbol{\alpha}}$$

<u>View #3</u>

There is a matrix, $\mathbf{V}^{-1}$, that maps vector **v** into vector α.

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

Now have a way to go back-and-forth between vector **v** and its coefficient vector α

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

Basis Matrix for ON Basis: We get a special structure!!!

Result: For an ON basis matrix **V**… $\mathbf{V}^{-1} = \mathbf{V}^{H}$

(the superscript H denotes "hermitian transpose", which consists of transposing the matrix and conjugating the elements)

To see this:

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

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

Inner products are 0 or 1 because this is an ON basis

Unitary and Orthogonal Matrices

A unitary matrix is a complex matrix $\mathbf{A}$ whose inverse is $\mathbf{A}^{-1} = \mathbf{A}^H$

For the real-valued matrix case… we get a special case of "unitary"
the idea of "unitary matrix" becomes "orthogonal matrix"
for which $\mathbf{A}^{-1} = \mathbf{A}^T$

Two Properties of Unitary Matrices: Let $\mathbf{U}$ be a unitary matrix and let $\mathbf{y}_1 = \mathbf{U}\mathbf{x}_1$ and $\mathbf{y}_2 = \mathbf{U}\mathbf{x}_2$

1. They preserve norms: $\|\mathbf{y}_i\| = \|\mathbf{x}_i\|$.
2. They preserve inner products: $< \mathbf{y}_1, \mathbf{y}_2 > = < \mathbf{x}_1, \mathbf{x}_2 >$

That is the "geometry" of the old space is preserved by the unitary matrix as it transforms into the new space.

(These are the same as the preservation properties of ON basis.)

DFT from Unitary Matrix Viewpoint:

Consider a discrete-time signal $x[n]$ for $n = 0, 1, \ldots N$-1.

We've already seen the DFT in a basis viewpoint: $\mathbf{x} = \sum_{k=0}^{N-1} \underbrace{\frac{1}{N} X[k]}_{\alpha_k} \mathbf{d}_k$

Now we can view the DFT as a transform from the Unitary matrix viewpoint:

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

DFT

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

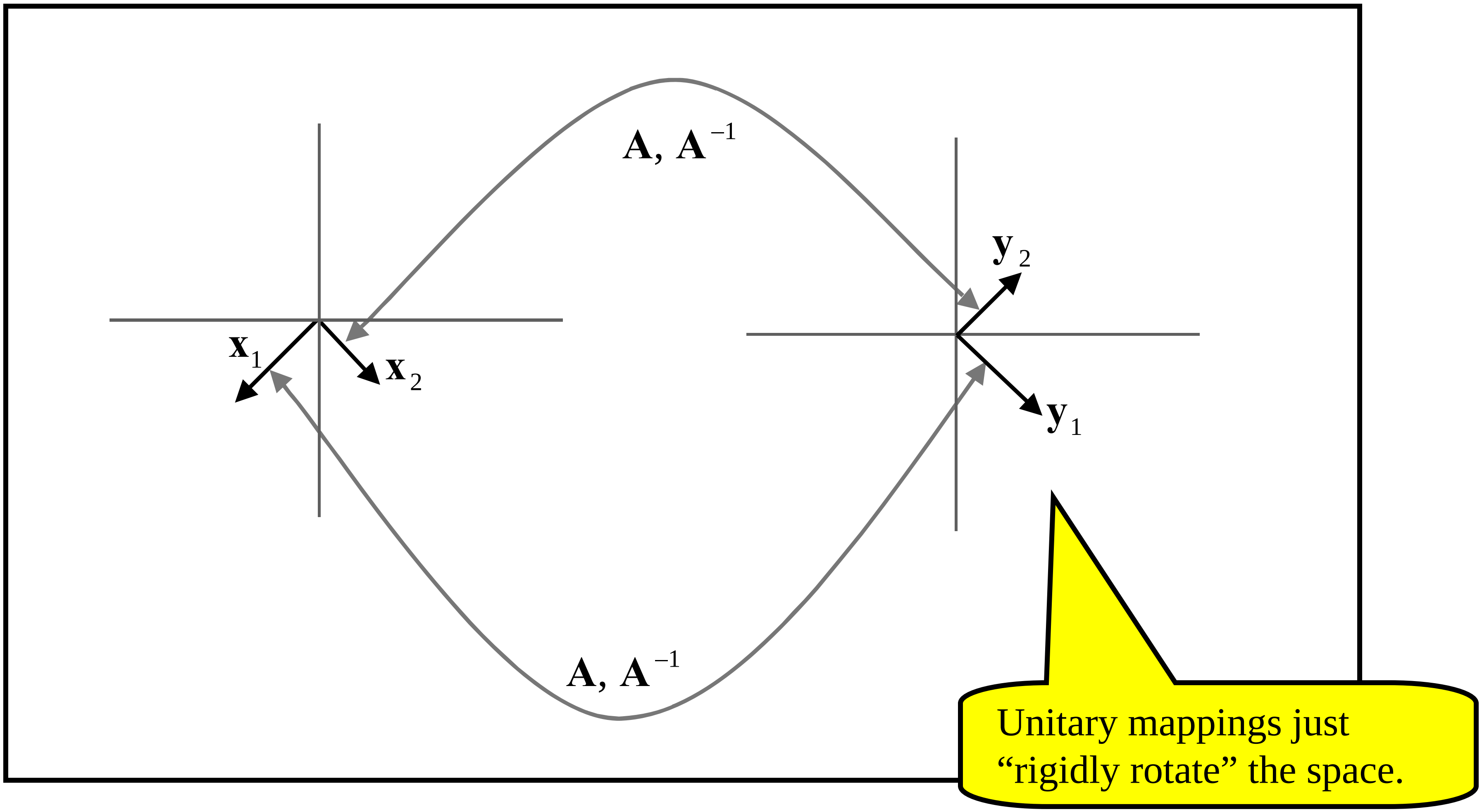
IDFT

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

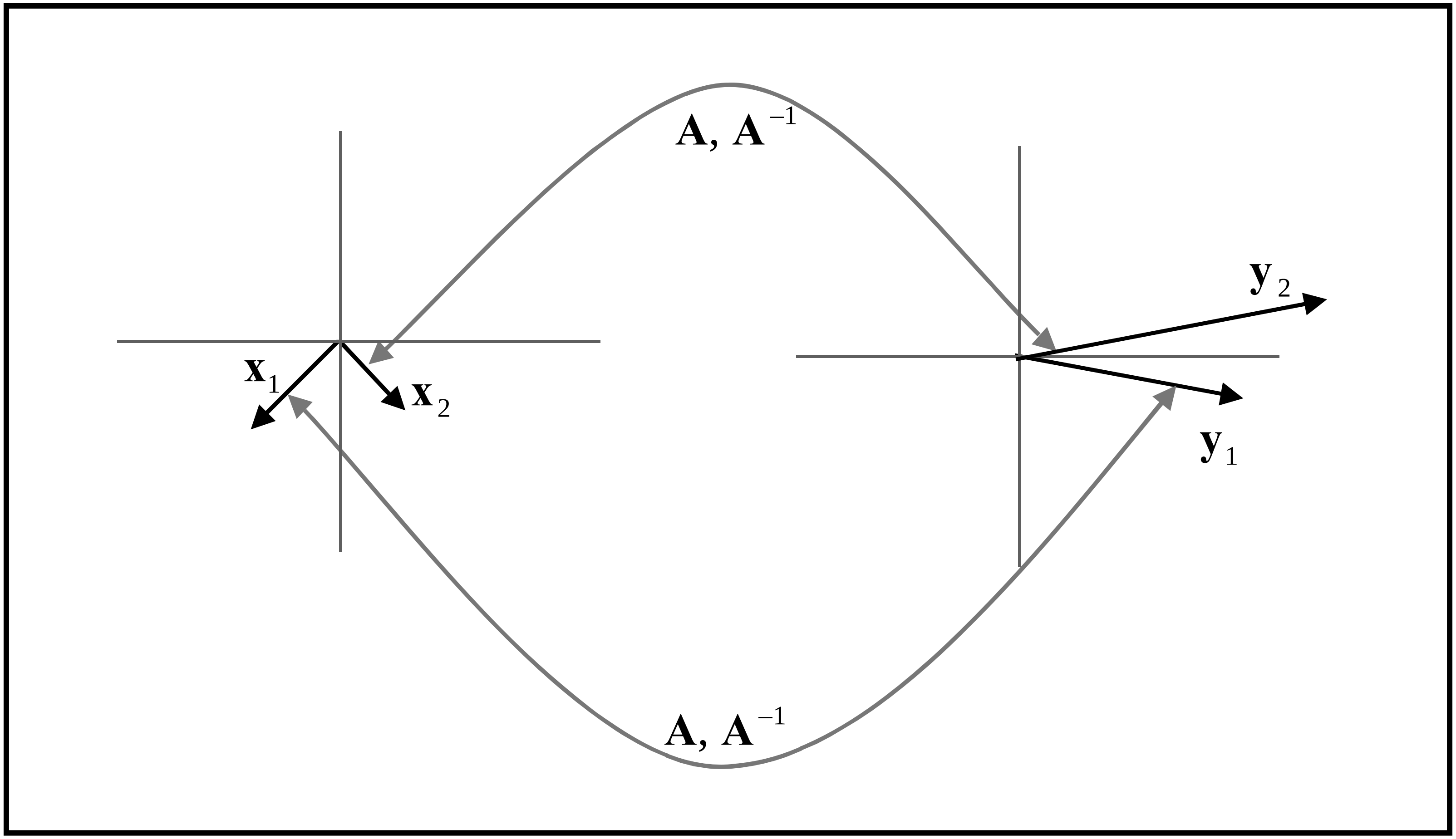
(*Acutally* **D** *is not unitary but* $N^{-1/2}$**D** *is unitary… see reading notes*)

Geometry Preservation of Unitary Matrix Mappings

Recall… unitary matrices map in such a way that the sizes of vectors and the orientation between vectors is not changed.



Effect of Non-Unitary Matrix Mappings



More on Matrices as Transforms

We'll limit ourselves here to real-valued vectors and matrices

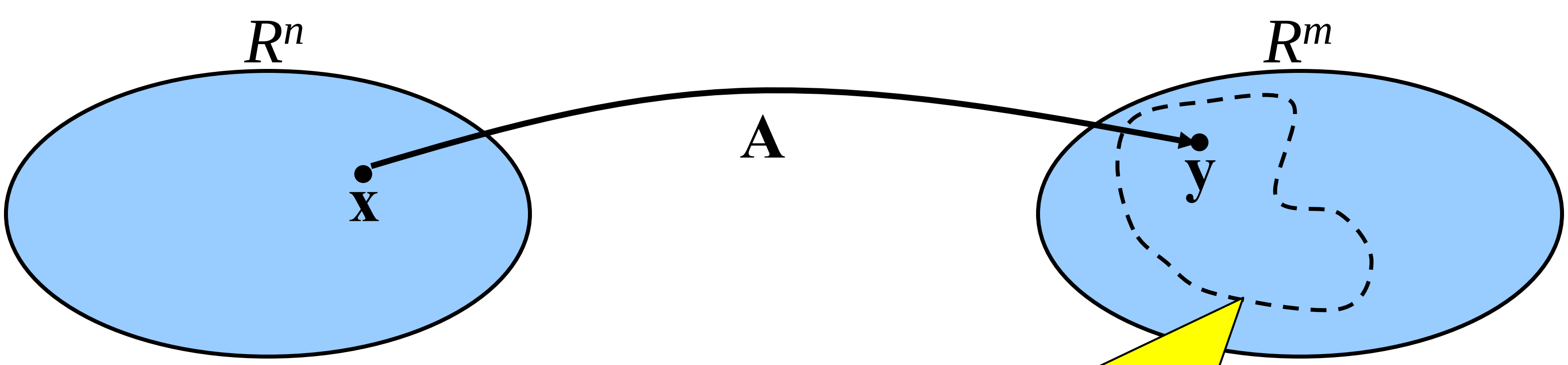
$$\underbrace{\mathbf{y}}_{m\times 1} = \underbrace{\mathbf{A}}_{m\times n}\underbrace{\mathbf{x}}_{n\times 1}$$

A maps <u>any</u> vector **x** in R^n

into <u>some</u> vector **y** in R^m

Mostly interested in two cases:

1. **"Tall Matrix"** $m > n$
2. **"Square Matrix"** $m = n$



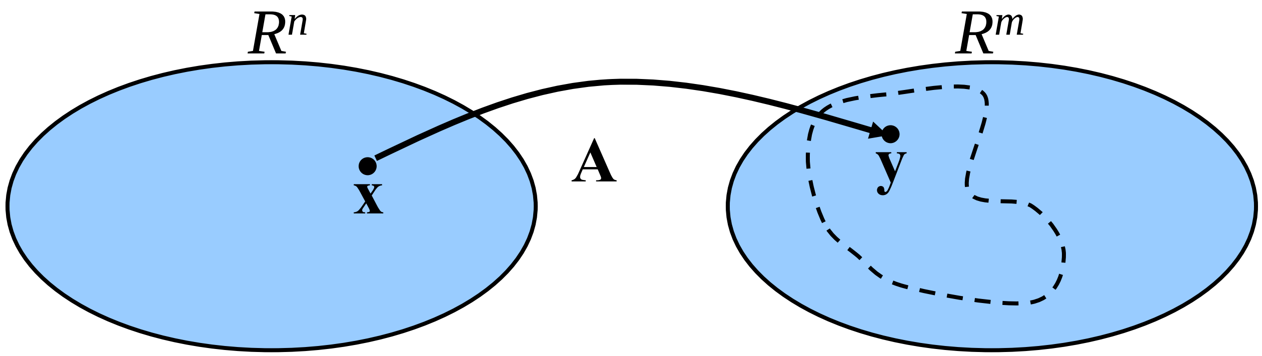
Range(**A**): "Range Space of **A**" = set of all vectors in R^m that can be reached by mapping

vector **y** = weighted sum of columns of **A**

⇒ may only be able to reach certain **y**'s

Range of a "Tall Matrix" ($m > n$)

The range($\mathbf{A}$) $\subset R^m$



"Proof": Since $\mathbf{y}$ is "built" from the n columns of $\mathbf{A}$ there are not enough to form a basis for R^m (they don't span R^m)

Range of a "Square Matrix" ($m = n$)

If the columns of A are linearly indep….The range($\mathbf{A}$) = R^m

…because the columns form a basis for R^m

Otherwise….The range($\mathbf{A}$) $\subset R^m$

…because the columns don't span R^m

<u>Rank of a Matrix</u>: rank(**A**) = largest # of linearly independent columns (or rows) of matrix **A**

For an $m \times n$ matrix we have that rank(**A**) $\leq$ min(m,n)

An $m \times n$ matrix **A** has "full rank" when rank(**A**) = min(m,n)

<u>Example</u>: This matrix has rank of 3 because the 4th column can be written as a combination of the first 3 columns

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

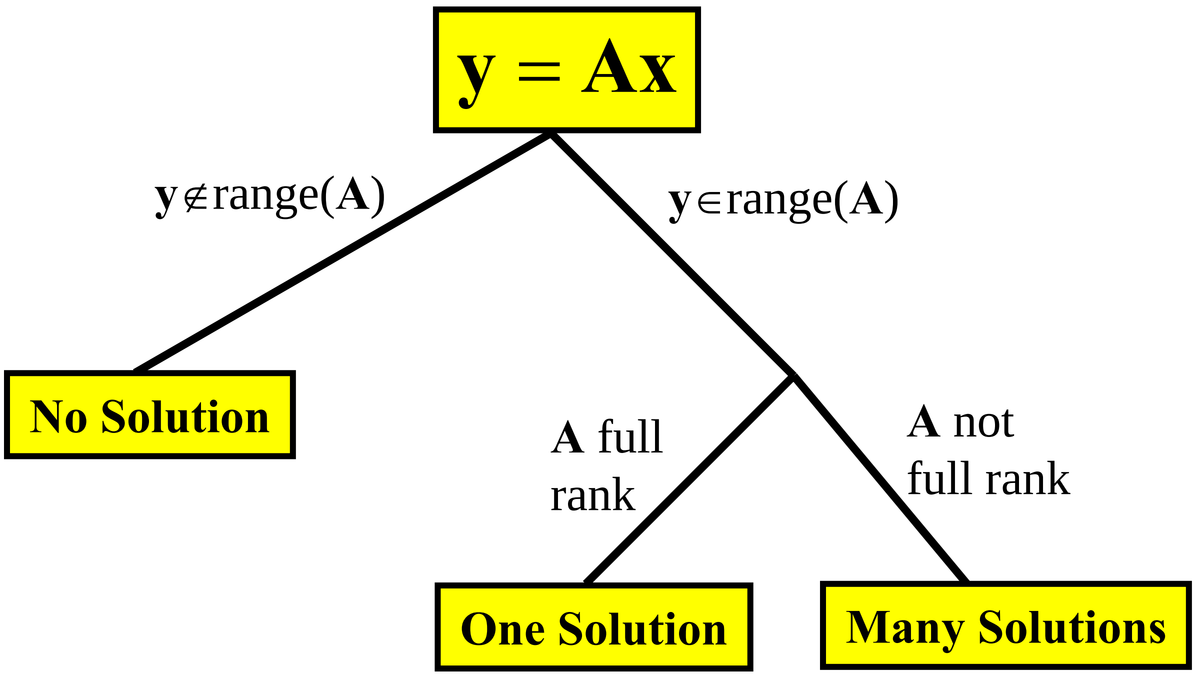
Characterizing “Tall Matrix” Mappings

We are interested in answering: Given a vector **y**, what vector **x** gets mapped into it via matrix **A**?

“Tall Matrix” ($m > n$) Case

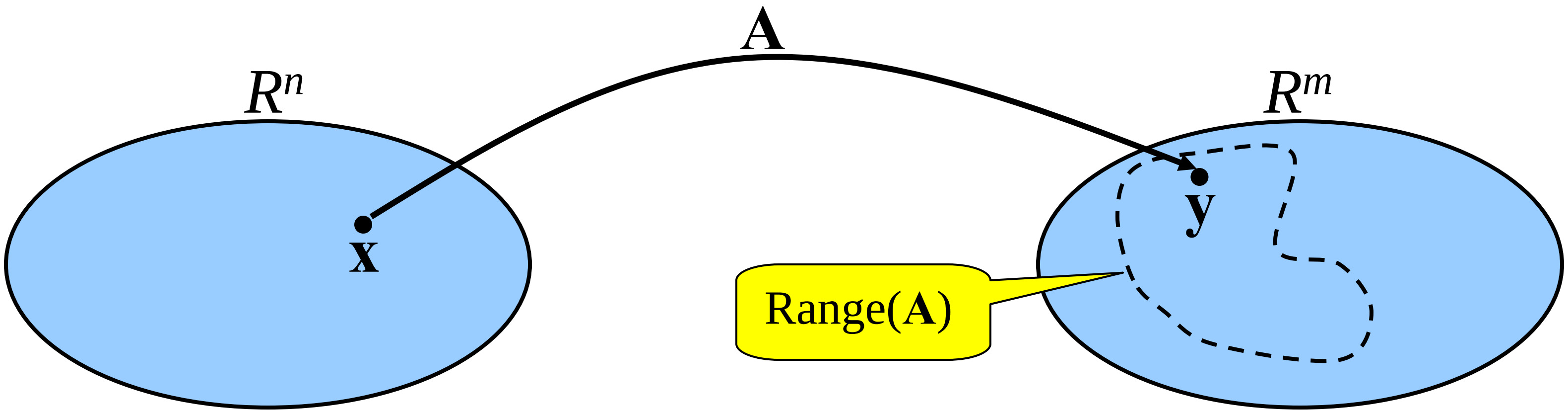
If **y** does not lie in range(**A**), then there is No Solution

If **y** lies in range(**A**), then there is a solution (but not necessarily just one unique solution)



Full-Rank "Tall Matrix" ($m > n$) Case

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



For a given $\mathbf{y} \in$range($\mathbf{A}$)…

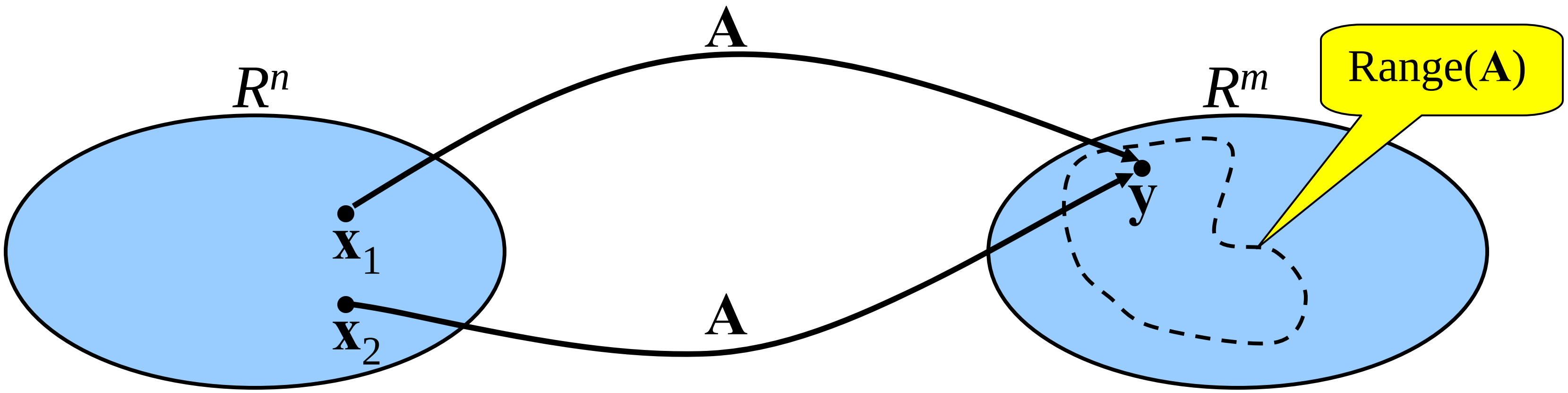
there is <u>only one</u> **x** that maps to it.

This is because the columns of **A** are linearly independent and we know from our studies of vector spaces that the coefficient vector of **y** is unique… **x** is that coefficient vector

By looking at **y** we can determine which **x** gave rise to it

<u>NonFull-Rank "Tall Matrix" ($m > n$) Case</u>

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



For a given $\mathbf{y} \in$range($\mathbf{A}$) there may be <u>more than one</u> $\mathbf{x}$ that maps to it

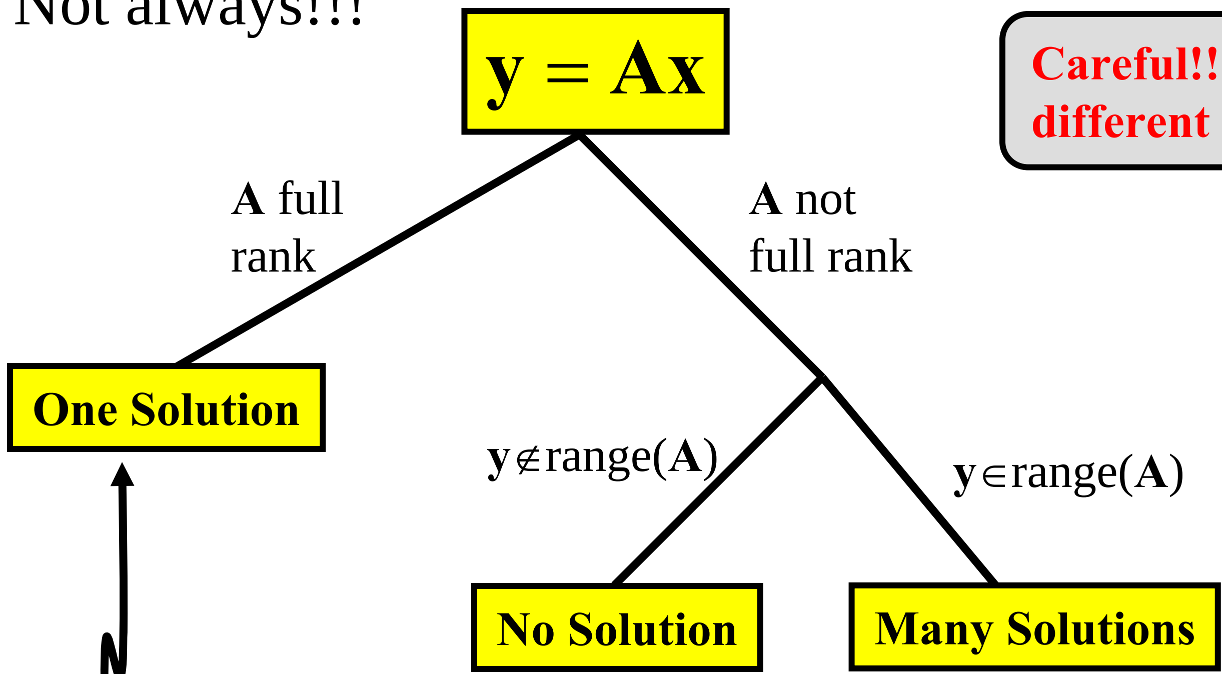
This is because the columns of $\mathbf{A}$ are linearly dependent and that redundancy provides several ways to combine them to create $\mathbf{y}$

By looking at $\mathbf{y}$ we can **<u>not</u>** determine which $\mathbf{x}$ gave rise to it

<u>Characterizing “Square Matrix” Mappings</u>

Q: Given <u>any</u> $\mathbf{y} \in R^n$ can we find an $\mathbf{x} \in R^n$ that maps to it?

A: Not always!!!



Careful!!! This is quite a different flow diagram here!!!

When a square **A** is full rank then its range covers the <u>complete</u> new space… then, **y** must be in range(**A**) and because the columns of **A** are a basis there is a way to build **y**

A Full-Rank Square Matrix is Invertible

A square matrix that has full rank is said to be….

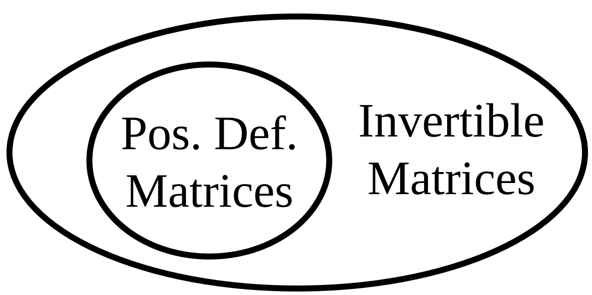
"nonsingular", "invertible"

Then we can find the **x** that mapped to **y** using $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$

Several ways to check if $n \times n$ **A** is invertible:

1. **A** is invertible if and only if (iff) its columns (or rows) are linearly independent (i.e., if it is full rank)
2. **A** is invertible iff $\det(\mathbf{A}) \neq 0$
3. **A** is invertible if (but not only if) it is "positive definite" (see later)
4. **A** is invertible iff all its eigenvalues are nonzero

   (because… det(A) = product of eigenvalues)



Eigenvalues and Eigenvectors of Square Matrices

If matrix **A** is $n \times n$, then **A** maps $R^n \rightarrow R^n$

Q: For a given $n \times n$ matrix **A**, which vectors get mapped into being <u>almost</u> themselves???

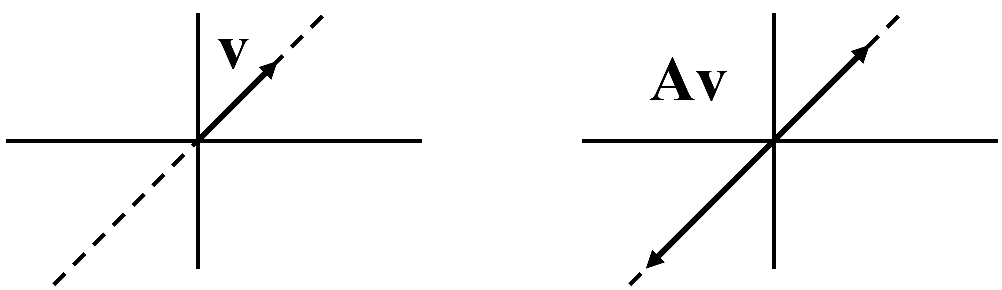
More precisely… Which vectors get mapped to a scalar multiple of themselves???

Even more precisely… which vectors **v** satisfy the following:

$$\boxed{\mathbf{A}\mathbf{v} = \lambda\mathbf{v}}$$

Input Output

These vectors are "special" and are called the eigenvectors of **A**. The scalar λ is that e-vector's corresponding eigenvalue.



"Eigen-Facts for Symmetric Matrices"

- If $n \times n$ real matrix $\mathbf{A}$ is symmetric, then
  - e-vectors corresponding to distinct e-values are orthonormal
  - e-values are real valued
  - can decompose $\mathbf{A}$ as

$$\boxed{\mathbf{A} = \mathbf{V\Lambda V}^T}$$

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix} \qquad \mathbf{VV}^T = \mathbf{I}$$

$$\mathbf{\Lambda} = diag\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$$

- If, further, $\mathbf{A}$ is pos. def. (semi-def.), then
  - e-values are positive (non-negative)
  - rank($\mathbf{A}$) = # of non-zero e-values
    - Pos. Def. $\Rightarrow$ Full Rank (and therefore invertible)
    - Pos. Semi-Def. $\Rightarrow$ Not Full Rank (and therefore not invertible)
  - When $\mathbf{A}$ is P. D., then we can write

$$\boxed{\mathbf{A}^{-1} = \mathbf{V\Lambda}^{-1}\mathbf{V}^T}$$

$$\mathbf{\Lambda}^{-1} = diag\left\{ 1/\lambda_1, 1/\lambda_2, \ldots, 1/\lambda_n \right\}$$

For P.D. $\mathbf{A}$, $\mathbf{A}^{-1}$ has the same e-vectors and has reciprocal e-values

Other Matrix Issues

We'll limit our discussion to real-valued matrices and vectors

Quadratic Forms and Positive-(Semi)Definite Matrices

Quadratic Form = Matrix form for a 2nd-order multivariate polynomial

Example:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

variable

fixed

The quadratic form of matrix **A** is:

$$Q_{\mathbf{A}}(x_1, x_2) = \mathbf{x}^T \mathbf{A} \mathbf{x} \qquad (1 \times 2) \cdot (2 \times 2) \cdot (2 \times 1) = (1 \times 1) \text{ scalar}$$

scalar

$$= \sum_{i=1}^{2} \sum_{j=1}^{2} a_{ij} x_i x_j = a_{11} x_1^2 + a_{22} x_2^2 + (a_{12} + a_{21}) x_1 x_2$$

- Values of the elements of matrix **A** determine the characteristics of the quadratic form $Q_{\mathbf{A}}(\mathbf{x})$
  - If $Q_{\mathbf{A}}(\mathbf{x}) \geq 0 \;\; \forall \mathbf{x} \neq \mathbf{0}$… then say that $Q_{\mathbf{A}}(\mathbf{x})$ is "<u>positive semi-definite</u>"
  - If $Q_{\mathbf{A}}(\mathbf{x}) > 0 \;\; \forall \mathbf{x} \neq \mathbf{0}$… then say that $Q_{\mathbf{A}}(\mathbf{x})$ is "<u>positive definite</u>"
  - Otherwise say that $Q_{\mathbf{A}}(\mathbf{x})$ is "<u>non-definite</u>"
- These terms carry over to the matrix that defines the Quad Form
  - If $Q_{\mathbf{A}}(\mathbf{x}) \geq 0 \;\; \forall \mathbf{x} \neq \mathbf{0}$… then say that **A** is "<u>positive semi-definite</u>"
  - If $Q_{\mathbf{A}}(\mathbf{x}) > 0 \;\; \forall \mathbf{x} \neq \mathbf{0}$… then say that **A** is "<u>positive definite</u>"