Advanced Transform Methods

Linear Algebra

Andy Watson

Linear Algebra

- Many Time-Frequency Transforms are linear, like matrices – so we need to know about vectors & matrices
- This lecture is a quick reminder of some of these concepts – called “Linear Algebra”
- Will keep to real-valued examples here

(For further reading, see for example,
W. G. Strang: Linear Algebra and its Applications (Academic Press, 1980)

Strang’s videotaped lecture course can be viewed at
http://web.mit.edu/18.06/www/Video/video-fall-99.html)

Scalars, Vectors and Matrices

- Scalar – completely determined by a single number.
  E.g. length, volume, brightness.
- Typical notation: lower case italic: a
- *Matlab example: Assign a single value:* » ***`a=5.3`***

- Vector – determined by magnitude and direction.
  Often given as several values (*elements* or *components*)
- *Dimensionality* – number of elements needed.
- Typical notation: lower case bold: $\mathbf{b}$
  (or alternatively underline: $\underline{b}$)
- *Matlab example:* » ***`b=[4; 3; 5]`***
- **This is a *column* vector**

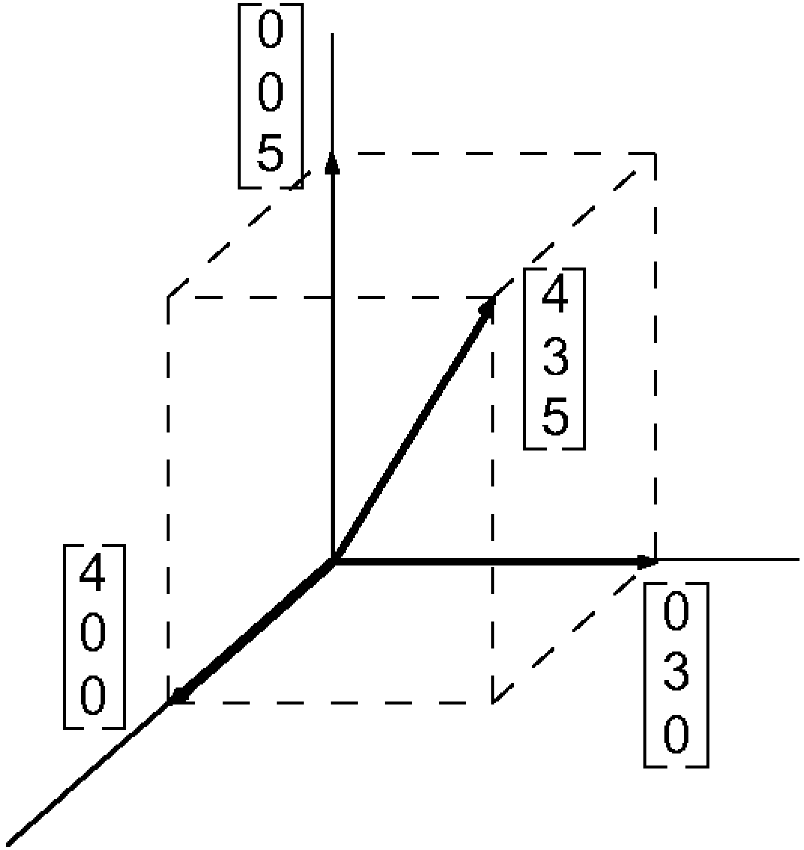
$$\text{Example}:\mathbf{b}=\begin{bmatrix}4\\3\\5\end{bmatrix}$$

Geometrical Representation

Components 4, 3, 5 are the coordinates of a point in 3-dimensional space.

Coordinates indicated by subscripts, e.g. $b_3 = 5$

Notation:
Scalar b_3 is a component of vector $\mathbf{b}$



Matrices

- Matrix – a rectangular array of scalars (its *elements*)
- Matrices said to be m x n ("m by n")
  m rows and n columns, with a total of mn elements.
- Notation: upper case bold: $\mathbf{A}$ (or double-underline: $\underline{\underline{A}}$)

$$\text{Example of } 3\times 4 \text{ matrix} : \mathbf{A} = \begin{bmatrix} 2 & 5 & 3 & 6 \\ 7 & 3 & 2 & 1 \\ 5 & 2 & 0 & 3 \end{bmatrix}$$

- *Matlab:* » `A=[2 5 3 6; 7 3 2 1; 5 2 0 3]`
- Elements indicated by subscripts: a_{ij}
  e.g. $a_{2,4} = 1$ Use comma if necessary: $a_{2,4}$ vs a_{24}

$$\text{Notation : Sometimes useful} : \mathbf{A} = [a_{ij}] \text{ or } a_{ij} = [\mathbf{A}]_{ij}$$

Transpose

- Transpose flips matrix along its diagonal: swaps rows and columns:

$$\text{If } \mathbf{A} = \begin{bmatrix} 2 & 5 & 3 & 6 \\ 7 & 3 & 2 & 1 \\ 5 & 2 & 0 & 3 \end{bmatrix} \text{ then transpose is } \mathbf{A}^t = \begin{bmatrix} 2 & 7 & 5 \\ 5 & 3 & 2 \\ 3 & 2 & 0 \\ 6 & 1 & 3 \end{bmatrix}$$

- Typical notation superscript t or T: $\mathbf{A}^t$ or $\mathbf{A}^T$
- *Matlab: use single quote » **A'***
- Column vector is 1 x n, transpose is n x 1 *row vector*:

$$\mathbf{b} = \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix} \text{ with transpose } \mathbf{b}^t = \begin{bmatrix} 4 & 3 & 5 \end{bmatrix}$$

If $\mathbf{A}^t = \mathbf{A}$, the matrix $\mathbf{A}$ is called *symmetric*.

Matrix and vector addition

- Vectors and matrices are added by adding elements

  Example: $\begin{bmatrix} 1 & 4 & 3 \\ 5 & 4 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 4 & 1 \\ 2 & 6 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 8 & 4 \\ 7 & 10 & 3 \end{bmatrix}$

  We write: $\mathbf{v} + \mathbf{w} = \mathbf{z}$ (vectors) or $\mathbf{A} + \mathbf{B} = \mathbf{C}$ (matrices)

  Expands to: $v_i + w_i = z_i$ for all i, and $a_{ij} + b_{ij} = c_{ij}$ for all i, j
- Matrices and vectors added must have same shape!
- *Matlab: » **A+B***

  Addition is commutative: $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$

  Addition is associative: $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$

  Scalar multiplication: $c\mathbf{A} = [ca_{ij}]$ (scales each element)
- *Matlab: » **c*A***

Inner Product ("dot product")

For two (real) vectors $\mathbf{a}$ and $\mathbf{b}$ of same dimension n

inner product is $\langle \mathbf{a}, \mathbf{b} \rangle = \sum_i a_i b_i$

Sometimes write as $\mathbf{a} \cdot \mathbf{b}$ ("dot product") or as $\mathbf{a}^t \mathbf{b}$

Example : given $\mathbf{a} = \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix}$ we get

$$\mathbf{a}^t \mathbf{b} = \begin{bmatrix} 2 & 5 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix} = 2 \times 4 + 5 \times 3 + 1 \times 5 = 28$$

- *Matlab: » **a'*b***

Norm (length) of a vector

Norm $\|\mathbf{a}\|$ of a vector $\mathbf{a}$ (strictly : its "2 - norm") is given by

$$\sqrt{\langle \mathbf{a}, \mathbf{a} \rangle} = \sqrt{\mathbf{a}^t \mathbf{a}} = \left(\sum_i a_i^2 \right)^{1/2}$$

Norm of a vector is commonly known as its *length*

- *Matlab:* » ***norm(a)*** *or* » ***sqrt(a'*a)***

Can show that $\|\langle \mathbf{a}, \mathbf{b} \rangle\| \leq \|\mathbf{a}\| \cdot \|\mathbf{b}\|$

Orthogonal and Orthonormal Vectors

Vectors $\mathbf{a}$ and $\mathbf{b}$ are *orthogonal* to each other if

$\mathbf{a}^t\mathbf{b} = 0$ i.e. their inner product $\langle \mathbf{a}, \mathbf{b} \rangle$ is zero.

Vectors are *orthonormal* if orthogonal and of unit length.

A *set* of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \quad , \mathbf{v}_m\}$, [i.e. $\{\mathbf{v}_i\}$ with $i = 1,2,...,m$] is orthonormal if

$$\mathbf{v}_i^t \mathbf{v}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{otherwise} \end{cases}$$

Example: The 3-dimensional basis system given by

$$\mathbf{v}_1 = [1\ 0\ 0],\ \mathbf{v}_2 = [0\ 1\ 0],\ \mathbf{v}_3 = [0\ 0\ 1]$$

is orthonormal.

For n-dimensional vectors, we can have no more than n vectors in any set of orthogonal vectors.

Matrix Multiplication

Two matrices $\mathbf{A}$ and $\mathbf{B}$ can be multiplied if the number of columns of $\mathbf{A}$ matches the number of rows of $\mathbf{B}$,
i.e. $\mathbf{A}$ is $m \times n$ and $\mathbf{B}$ is $n \times p$ for some integers m, n, p.
Element ik of product $\mathbf{AB}$ is: $$[\mathbf{AB}]_{ik} = \sum_j a_{ij} b_{jk}$$

[Or : in Einstein's summation convention : $[\mathbf{AB}]_{ik} = a_{ij} b_{jk}$]

Sum repeated Indices (here: j)

Matrix multiplication is

Associative: $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$

Distributive: $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$

but not Commutative: $\mathbf{AB} \neq \mathbf{BA}$ (except in special cases)

Useful identities:

$$\left(\mathbf{A}^t\right)^t = \mathbf{A} \qquad (\mathbf{A} + \mathbf{B})^t = \mathbf{A}^t + \mathbf{B}^t \qquad (\mathbf{AB})^t = \mathbf{B}^t \mathbf{A}^t$$

Special Matrices

Diagonal matrix : non - zero elements only occur on diagonal.

Example : $\mathbf{D} = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 7 & 0 \end{bmatrix}$ Note $d_{ij} = 0$ if $i \neq j$

Diagonal matrices are often square.

Matlab : **`diag(.)`** gets diagonal elements or makes matrix

Identity matrix $\mathbf{I}$: square diagonal matrix with ones on diagonal.

Example for 3x3 : $\mathbf{I}_3 = \mathbf{I}_{3\times 3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Multiplication by identity matrix leaves a matrix unchanged :

$$\mathbf{A}_{m\times n}\mathbf{I}_{n\times n} = \mathbf{A}_{m\times n} = \mathbf{I}_{m\times m}\mathbf{A}_{m\times n}$$

Matrix Inverse

Inverse $\mathbf{A}^{-1}$ of a matrix $\mathbf{A}$ is the matrix where $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$

For $\mathbf{A}^{-1}$ to exist, $\mathbf{A}$ must be square.

If $\mathbf{A}$ has an inverse, it is said to be *invertible*, and $\mathbf{A}^{-1}$ is unique.

The product of invertible matrices has an inverse:

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \qquad (\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$$

The identity matrix is its own inverse: $\mathbf{I}^{-1} = \mathbf{I}$

(Methods exist for calculating inverses)

Matlab: **`inv(A)`** calculates the inverse of a matrix.

Linear Independence

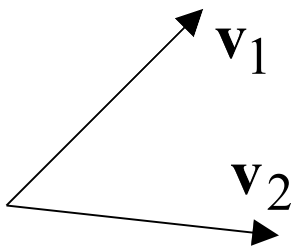
Given a set of vectors $\{\mathbf{v}_i\}\ \ i = 1,2,...,k$ if we can find a set of scalars $\{c_i\}\ \ i = 1,2,...,k$ (excluding $c_1 = c_2 = \quad = 0$) such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \quad + c_k\mathbf{v}_k = \mathbf{0}$$

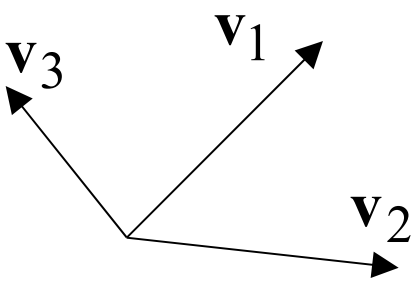
then the set of vectors is called *linearly dependent*.

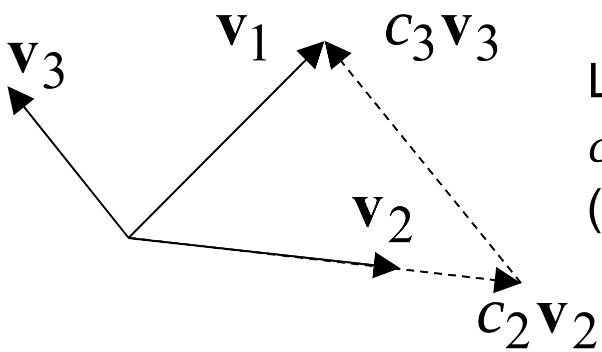
If there is none, they are called *linearly independent*.

Visualialization:
2-dimensional plane



Linearly independent





Linearly *dependent*

$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = 0$

(using $c_3 = -1$)

Max n linearly independent vectors in n-dimensional space

Spanning a vector space

The real vector space R^n is the space of vectors which each have n real components. The exponent n is called the *dimension* of the space.

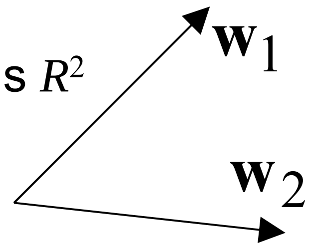
E.g. R^3 is our "usual" three-dimensional space.

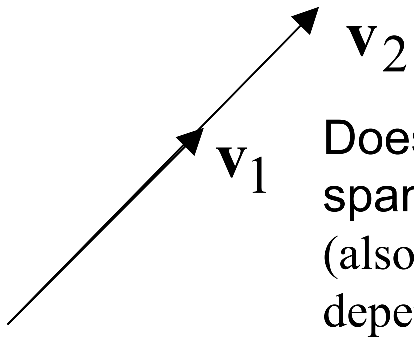
Given an n-dimensional real vector space R^n, a set of vectors $\{\mathbf{w}_i\}\ \ i = 1,2,...,k$ is said to *span* the space if every vector $\mathbf{v}$ in R^n can be written as

$$\mathbf{v} = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \quad + c_k\mathbf{w}_k$$

for some set of scalars $\{c_i\}\ \ i = 1,2,...,k$

Spans R^2





Does not span R^2 (also linearly dependent)

Basis vectors

A set of vectors $\{\mathbf{w}_i\}$ $i = 1, 2, \ldots, n$ is said to be a *basis* for a vector space R^n if the set of vectors $\{\mathbf{w}_i\}$

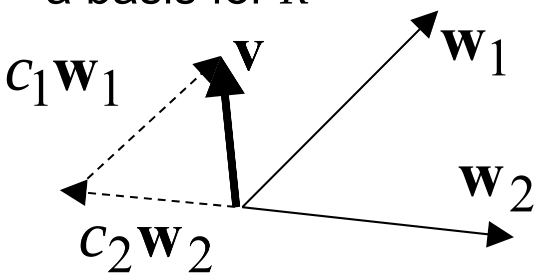
(a) is linearly independent, and

(b) spans the space R^n.

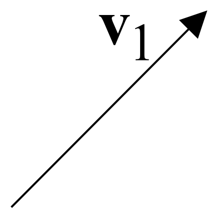
There must be exactly n vectors in the basis.

Each vector $\mathbf{v}$ in R^n has a unique set of *coordinates* $\{c_i\}$ in the basis $\{\mathbf{w}_i\}$, where $\mathbf{v} = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \quad + c_n\mathbf{w}_n$

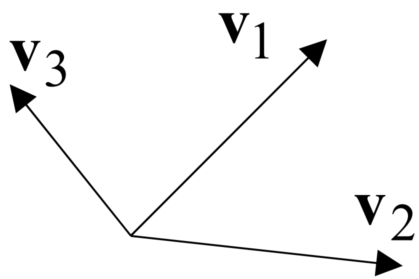
Set $\{\mathbf{w}_1, \mathbf{w}_2\}$ is a basis for R^2



Not a basis for R^2
(Doesn't span R^2)



Not a basis for R^2
(linearly dependent)



Row Space & Column Space; Rank

The column space $R(\mathbf{A})$: space spanned by the columns of $\mathbf{A}$.

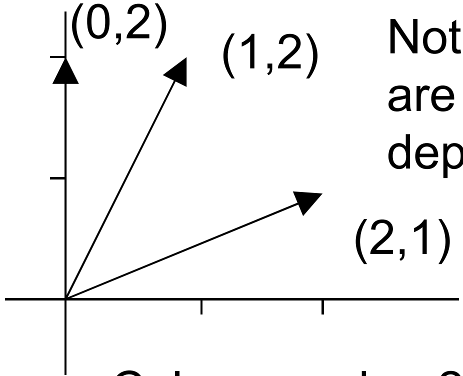
The row space $R(\mathbf{A}^t)$ is the space spanned by the rows of $\mathbf{A}$.

The row space and column space have the same dimensionality r, the *rank* of $\mathbf{A}$.

This fact sometimes expressed as "row rank $=$ column rank".

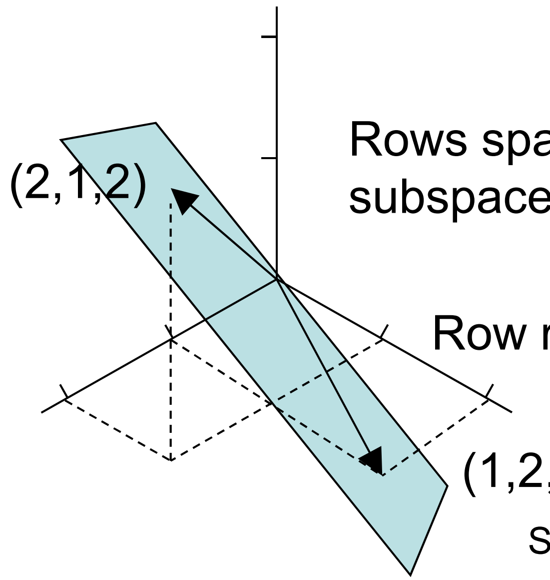
Example : $\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 2 \end{bmatrix}$

Columns span entire R^2



Note that columns are linearly dependent

Column rank = 2



Rows span 2-d subspace of R^3

Row rank = 2

Null Space; Orthogonal Subspaces

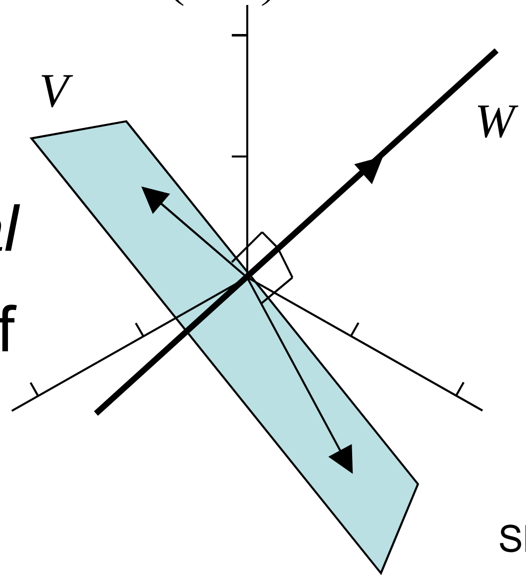
The *nullspace* $N(\mathbf{A})$ of $\mathbf{A}$ is the space *not* spanned by the rows of $\mathbf{A}$. This has dimension $n-r$.

Two subspaces V and W are orthogonal if every vector $\mathbf{v}$ in V is orthogonal to every vector $\mathbf{w}$ in W.

I.e. we must have $\mathbf{v}^t\mathbf{w}=0$ for all $\mathbf{v}\in V, \mathbf{w}\in W$.

So, the nullspace $N(\mathbf{A})$ and row space $R(\mathbf{A}^t)$ are *orthogonal*.

Example: 2-d subspace V (plane) is orthogonal to 1-d subspace W (line)



In the diagram, W is the *orthogonal complement* $V^\perp$ of V (the space of all vectors orthogonal to V).

Bases and Hilbert Spaces

- A set of vectors or functions $\{\Psi_n\}$ spans a vector space if any element of that space can be expressed as a linear combination of members of that set

$$s = \sum_n c_n \Psi_n$$

- $\{\Psi_n\}$ is a basis set if the c_n are unique.
- The set is an orthogonal basis if $n \neq m \Rightarrow \langle \Psi_n, \Psi_m \rangle = 0$
- The set is an orthonormal basis if $n = m \Rightarrow \langle \Psi_n, \Psi_m \rangle = 1$
- A space which has all these properties is a type of Hilbert space
- Real numbers, complex numbers and integrable functions form Hilbert Spaces

Properties of Orthonormal Bases

- If $\{\psi_n\}$ constitutes a basis for V, then any vector or function in V can be written as

$$s = \sum_n c_n \Psi_n$$

- However, c_n may be difficult to compute. If $\{\psi_n\}$ form an orthonormal basis, this difficulty is eliminated, since then

$$c_n = \langle s, \Psi_n \rangle$$

- Thus if $\{\psi_n\}$ is a set of orthonormal basis for V, then any ***s*** in V can be written as

$$s = \sum_j \langle s, \Psi_j \rangle \Psi_j$$

$$= \langle s, \Psi_1 \rangle \Psi_1 + \langle s, \Psi_2 \rangle \Psi_2 + \ldots\ldots + \langle s, \Psi_n \rangle \Psi_n$$

Biorthogonal Bases

- Sometimes, orthonormal bases are not available. A generalization is the concept of **biorthogonal bases ,** which are actually a pair of bases (that are linearly independent).
- If $\{\psi_n\}$ and $\{\hat{\Psi}_n\}$ are both basis vectors themselves for *V*, and satisfy

$$\left\langle \Psi_i , \hat{\Psi}_j \right\rangle = \delta_{ij}$$

then any **s** in *V* can be written as

$$s = \sum_{j=1}^{n} \left\langle s , \Psi_j \right\rangle \hat{\Psi}_j$$

Dual Base example

- A dual basis may be linearly dependent, a biorthogonal one may not
- **A basis set $\{\hat{\Psi}_i\}$ is said to be the dual basis of $\{\Psi_i\}$ if the** biorthogonality **condition**

$$\left\langle \Psi_i, \hat{\Psi}_j \right\rangle = \sum_k \Psi_i(k)\hat{\Psi}_j(k) = \delta_{ij}$$

is satisfied. (This is the Kronecker Delta Function.)

- **An example of a dual basis**

$$\left\{\Psi_i\right\} = \left\{(1,0),(1,1)\right\}$$

$$\left\{\hat{\Psi}_i\right\} = \left\{(1,-1),(0,1)\right\}$$

Frames

- Sometimes, biorthogonal bases are not available either. We'd like to represent any vector in V as a linear combination of some simpler vectors / functions, while giving up orthogonality and even linear independence:

$$s = \sum_{j=1}^{n} \left\langle s, \Psi_j \right\rangle \hat{\Psi}_j = \sum_{j=1}^{n} \left\langle s, \hat{\Psi}_j \right\rangle \Psi_j$$

- Vectors $\{\psi_n\}$ form a **frame** with frame bounds *A*, *B* if for any vector *s* in V

$$A\|s\|^2 \le \sum_j \left| \left\langle s, \Psi_j \right\rangle \right|^2 \le B\|s\|^2$$

Frames

- If $A=B$, the frame is **tight**. If removing an element from the frame makes it no longer a frame, then the original frame is said to be **exact.**
- If the original frame is a basis, then the two frames form a **biorthogonal basis system**
- Any vector can be written as a linear combination of frame vectors. However, coefficients are no longer unique, vectors no longer independent, and the frame does not constitute a basis. Only exact frames are bases.

Tight Frame example

- In finite dimensions, vectors can always be removed from a frame to get a basis, but in infinite dimensions, that is not always possible.
- Example of frame in finite dimensions is a matrix more rows than columns but with independent rows.
- Example of tight frame is a matrix with more rows than columns but with orthogonal rows.
- A 3d tight frame in 2d space
  - a family of three vectors in the plane which are obtained by successive rotations of a third of turn of one vector

$$\Psi_1 = (0,1)$$

$$\Psi_2 = (-\sqrt{3/4}, -1/2)$$

$$\Psi_3 = (\sqrt{3/4,}\ -1/2)$$