Machine Learning Mathematics - Advance Linear Algebra

Machine Learning II

Lecture 2-c

THE GEORGE WASHINGTON UNIVERSITY

WASHINGTON, DC

Notation

Vectors in $\Re^n$.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

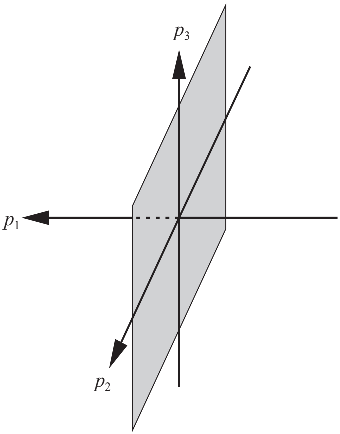
Generalized Vectors.

$$\mathcal{X}$$

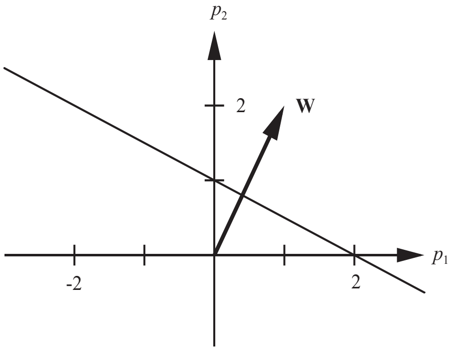
1. An operation called vector addition is defined such that if $x \in X$ and $y \in X$ then $x + y \in X$.
2. $x + y = y + x$
3. $(x + y) + z = x + (y + z)$
4. There is a unique vector $0 \in X$, called the zero vector, such that $x + 0 = x$ for all $x \in X$.
5. For each vector there is a unique vector in X, to be called $(-x)$, such that $x + (-x) = 0$.

6. An operation, called multiplication, is defined such that for all scalars $a \in F$, and all vectors $x \in X$, $a\,x \in X$.
7. For any $x \in X$, $1x = x$ (for scalar 1).
8. For any two scalars $a \in F$ and $b \in F$, and any $x \in X$, $a(bx) = (ab)\,x$.
9. $(a+b)\,x = a\,x + b\,x$.
10. $a(x+y) = a\,x + a\,y$

Is the p_2, p_3 plane a vector space?



Is the line $p_1 + 2p_2 - 2 = 0$ a vector space?

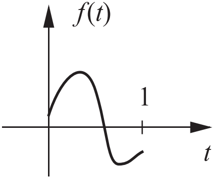


Polynomials of degree 2 or less.

$$x = 2 + t + 4t^2$$

$$y = 1 + 5t$$

Continuous functions in the interval [0,1].



If

$$a_1 \chi_1 + a_2 \chi_2 + \ldots + a_n \chi_n = \mathit{0}$$

implies that each

$$a_i = 0$$

then

$$\{\chi_i\}$$

is a set of linearly independent vectors.

$$\mathbf{p}_1 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \qquad \mathbf{p}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

Let

$$a_1\mathbf{p}_1 + a_2\mathbf{p}_2 = \mathbf{0}$$

$$\begin{bmatrix} -a_1 + a_2 \\ a_1 + a_2 \\ -a_1 + (-a_2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

This can only be true if

$$a_1 = a_2 = 0$$

Therefore the vectors are independent.

A subset **spans** a space if every vector in the space can be written as a linear combination of the vectors in the subspace.

$$x = x_1 u_1 + x_2 u_2 + \ldots + x_m u_m$$

- A set of basis vectors for the space X is a set of vectors which spans X and is linearly independent.
- The dimension of a vector space, Dim(X), is equal to the number of vectors in the basis set.
- Let X be a finite dimensional vector space, then every basis set of X has the same number of elements.

Polynomials of degree 2 or less.

Basis A:

$u_1 = 1 \qquad u_2 = t \qquad u_3 = t^2$

Basis B:

$u_1 = 1 - t \qquad u_2 = 1 + t \qquad u_3 = 1 + t + t^2$

(Any three linearly independent vectors in the space will work.)

How can you represent the vector $x = 1 + 2t$ using both basis sets?

A scalar function of vectors x and y can be defined as an **inner product**, (x,y), provided the following are satisfied (for real inner products):

- $(x,y) = (y,x)$
- $(x, ay_1 + by_2) = a(x,y_1) + b(x,y_2)$
- $(x,x) \geq 0$, where equality holds iff $x = 0$.

A scalar function of a vector x is called a **norm**, $\|x\|$, provided the following are satisfied:

- $\|x\| \geq 0$.
- $\|x\| = 0$ iff $x = 0$.
- $\|ax\| = |a|\,\|x\|$ for scalar a.
- $\|x + y\| \leq \|x\| + \|y\|$.

Standard Euclidean Inner Product

$$\mathbf{x}^{\mathrm{T}}\mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

Standard Euclidean Norm

$$\|x\| = (x, x)^{1/2}$$

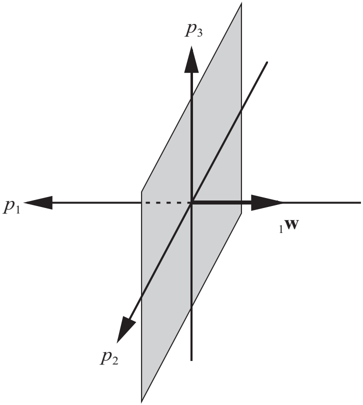
$$\|\mathbf{x}\| = (\mathbf{x}^T\mathbf{x})^{1/2} = (x_1{}^2 + x_2{}^2 + \ldots + x_n{}^2)^{1/2}$$

Angle

$$\cos(\theta) = (x,y)/(\|x\|\ \|y\|)$$

Two vectors $x, y \in X$ are orthogonal if $(x, y) = 0$.

Example



Any vector in the p_2,p_3 plane is orthogonal to the weight vector.

Independent Vectors

$y_1, y_2, \ldots, y_n$

Orthogonal Vectors

$v_1, v_2, \ldots, v_n$

Step 1: Set first orthogonal vector to first independent vector.

$$v_1 = y_1$$

Step 2: Subtract the portion of y_2 that is in the direction of v_1.

$$v_2 = y_2 - a v_1$$

Where a is chosen so that v_2 is orthogonal to v_1:

$$(v_1, v_2) = (v_1, y_2 - a v_1) = (v_1, y_2) - a(v_1, v_1) = 0$$

$$a = \frac{(v_1, y_2)}{(v_1, v_1)}$$

Projection of y_2 on v_1:

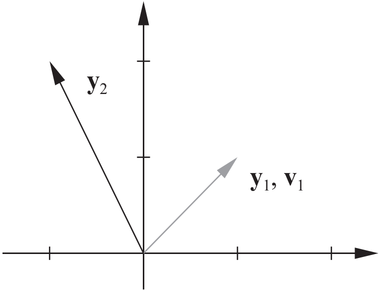
$$\frac{(v_1,y_2)}{(v_1,v_1)}\ v_1$$

Step k: Subtract the portion of y_k that is in the direction of all previous v_i.

$$v_k = y_k - \sum_{i=1}^{k-1} \frac{(v_i,y_k)}{(v_i,v_i)}\ v_i$$

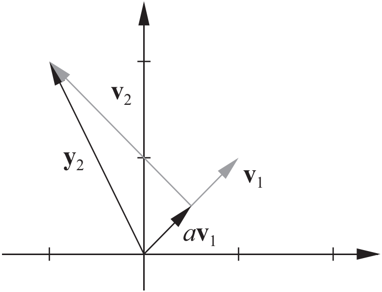
$$\mathbf{y}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \mathbf{y}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

Step 1. $\quad \mathbf{v}_1 = \mathbf{y}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$



Step 2.

$$\mathbf{v}_2 = \mathbf{y}_2 - \frac{\mathbf{v}_1^T \mathbf{y}_2}{\mathbf{v}_1^T \mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix} - \frac{\begin{bmatrix} 1 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 2 \end{bmatrix}}{\begin{bmatrix} 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} - \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} -1.5 \\ 1.5 \end{bmatrix}$$



If a vector space X has a basis set $\{v_1, v_2, ..., v_n\}$, then any $\chi \in X$ has a unique vector expansion:

$$\chi = \sum_{i=1}^{n} x_i v_i = x_1 v_1 + x_2 v_2 + ... + x_n v_n$$

If the basis vectors are **orthogonal**, and we take the inner product of v_j and χ:

$$(v_j, \chi) = (v_j, \sum_{i=1}^{n} x_i v_i) = \sum_{i=1}^{n} x_j (v_j, v_i) = x_j (v_j, v_j)$$

Therefore the coefficients of the expansion can be computed:

$$x_j = \frac{(v_j, \chi)}{(v_j, v_j)}$$

The vector expansion provides a meaning for writing a vector as a column of numbers.

$$\chi = \sum_{i=1}^{n} x_i v_i = x_1 v_1 + x_2 v_2 + \ldots + x_n v_n$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

To interpret **x**, we need to know what basis was used for the expansion.

Definition of reciprocal basis vectors, r_i:

$$(r_i, v_j) = 0 \qquad i \neq j$$
$$= 1 \qquad i = j$$

where the basis vectors are $\{v_1, v_2, ..., v_n\}$, and the reciprocal basis vectors are $\{r_1, r_2, ..., r_n\}$.

For vectors in $\Re^n$ we can use the following inner product:

$$(r_i, v_j) = \mathbf{r}_i^T \mathbf{v}_j$$

Therefore, the equations for the reciprocal basis vectors become:

$$\mathbf{R}^T \mathbf{B} = \mathbf{I} \quad \Longrightarrow \quad \mathbf{R}^T = \mathbf{B}^{-1}$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{bmatrix} \qquad \mathbf{R} = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \dots & \mathbf{r}_n \end{bmatrix}$$

$$\chi = x_1 v_1 + x_2 v_2 + \ldots + x_n v_n$$

Take the inner product of the j th reciprocal basis vector with the vector to be expanded:

$$(r_j, \chi) = (r_j, \sum_{i=1}^{n} x_i v_i) = \sum_{i=1}^{n} x_j (r_j, v_i) = x_j (r_j, v_j) = x_j$$

Because, by definition of the reciprocal basis vectors:

$$\begin{aligned} (r_i, v_j) &= 0 \qquad i \neq j \\ &= 1 \qquad i = j \end{aligned}$$

In general, we then have (even for nonorthogonal basis vectors):

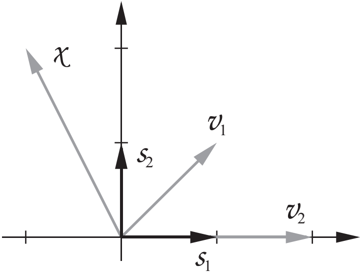
$$x_j = (r_j, \chi)$$

Basis Vectors:

$$\mathbf{v}_1^s = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2^s = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

Vector to Expand:

$$\mathbf{x}^s = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$



Reciprocal Basis Vectors:

$$\mathbf{R}^T = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 \\ 0.5 & -0.5 \end{bmatrix} \qquad \mathbf{r}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \mathbf{r}_2 = \begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix}$$

Expansion Coefficients:

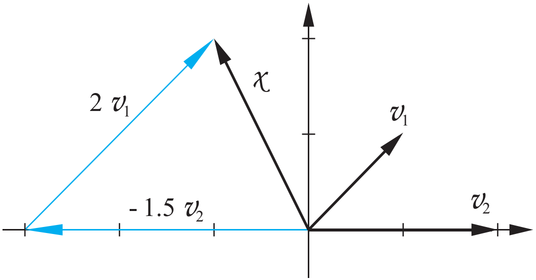
$$x_1^v = \mathbf{r}_1^T \mathbf{x}^s = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = 2$$

$$x_2^v = \mathbf{r}_2^T \mathbf{x}^s = \begin{bmatrix} 0.5 & -0.5 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = -1.5$$

Matrix Form:

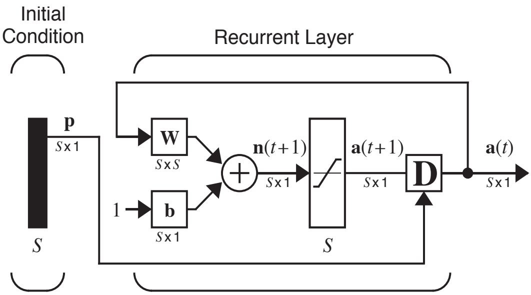
$$\mathbf{x}^v = \mathbf{R}^T \mathbf{x}^s = \mathbf{B}^{-1} \mathbf{x}^s = \begin{bmatrix} 0 & 1 \\ 0.5 & -0.5 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1.5 \end{bmatrix}$$

$$\chi = (-1)s_1 + 2s_2 = 2v_1 - 1.5v_2$$



$$\mathbf{x}^s = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \qquad \mathbf{x}^v = \begin{bmatrix} 2 \\ -1.5 \end{bmatrix}$$

The interpretation of the column of numbers depends on the basis set used for the expansion.



$$\mathbf{a}(0) = \mathbf{p} \qquad \mathbf{a}(t+1) = \mathbf{satlins}(\mathbf{Wa}(t)+\mathbf{b})$$

- The network output is repeatedly multiplied by the weight matrix W.
- What is the effect of this repeated operation?
- Will the output converge, go to infinity, oscillate?
- In this chapter we want to investigate matrix multiplication, which represents a general linear transformation.

A **transformation** consists of three parts:

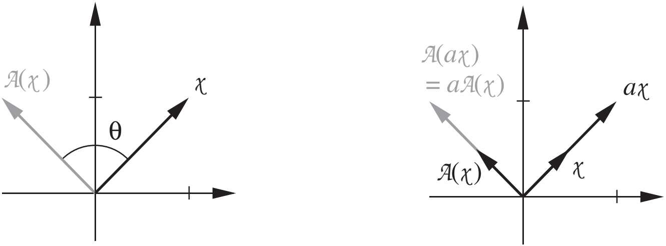
1. A set of elements $X = \{\chi_i\}$, called the domain,
2. A set of elements $Y = \{y_i\}$, called the range, and
3. A rule relating each $\chi_i \in X$ to an element $y_i \in Y$.

A transformation is **linear** if:

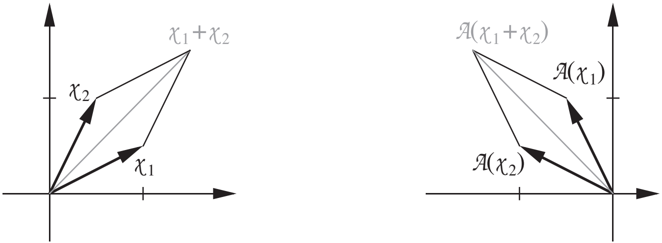
1. For all $\chi_1, \chi_2 \in X$, $\mathcal{A}(\chi_1 + \chi_2) = \mathcal{A}(\chi_1) + \mathcal{A}(\chi_2)$,
2. For all $\chi \in X$, $a \in \Re$, $\mathcal{A}(a\,\chi) = a\,\mathcal{A}(\chi)$.

Is rotation linear?

1.



2.



Any linear transformation between two finite-dimensional vector spaces can be represented by matrix multiplication.

Let $\{v_1, v_2, ..., v_n\}$ be a basis for X, and let $\{u_1, u_2, ..., u_m\}$ be a basis for Y.

$$x = \sum_{i=1}^{n} x_i v_i \qquad y = \sum_{i=1}^{m} y_i u_i$$

Let $\mathcal{A}: X \rightarrow Y$

$$\mathcal{A}(x) = y$$

$$\mathcal{A}\left(\sum_{j=1}^{n} x_j v_j\right) = \sum_{i=1}^{m} y_i u_i$$

Since A is a linear operator,

$$\sum_{j=1}^{n} x_j \mathcal{A}(v_j) = \sum_{i=1}^{m} y_i u_i$$

Since the u_i are a basis for Y,

$$\mathcal{A}(v_j) = \sum_{i=1}^{m} a_{ij} u_i$$

(The coefficients a_{ij} will make up the matrix representation of the transformation.)

$$\sum_{j=1}^{n} x_\text{j} \sum_{i=1}^{m} a_{ij} u_i = \sum_{i=1}^{m} y_i u_i$$

Matrix Representation 3

$$\sum_{i=1}^{m} u_i \sum_{j=1}^{n} a_{ij} x_j = \sum_{i=1}^{m} y_i u_i$$

$$\sum_{i=1}^{m} u_i \left(\sum_{j=1}^{n} a_{ij} x_j - y_i \right) = 0$$

Because the u_{i} are independent,

$$\sum_{j=1}^{n} a_{ij} x_j = y_i$$

This is equivalent to matrix multiplication.

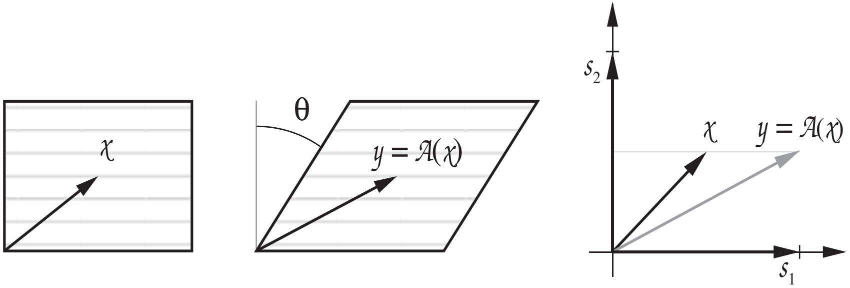
$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

- A linear transformation can be represented by matrix multiplication.
- To find the matrix which represents the transformation we must transform each basis vector for the domain and then expand the result in terms of the basis vectors of the range.

$$\boxed{\mathcal{A}(v_j) = \sum_{i=1}^{m} a_{ij} u_i}$$

Each of these equations gives us one column of the matrix.

Stand a deck of playing cards on edge so that you are looking at the deck sideways. Draw a vector x on the edge of the deck. Now "skew" the deck by an angle θ, as shown below, and note the new vector $y = A(x)$. What is the matrix of this transformation in terms of the standard basis set?



To find the matrix we need to transform each of the basis vectors.

$$\mathcal{A}(v_j) = \sum_{i=1}^{m} a_{ij} u_i$$

We will use the standard basis vectors for both the domain and the range.

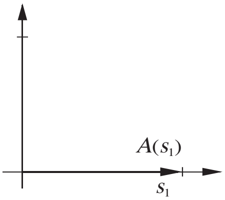
$$A(s_j) = \sum_{i=1}^{2} a_{ij} s_i = a_{1j} s_1 + a_{2j} s_2$$

6

Example - (3)

We begin with s_1:

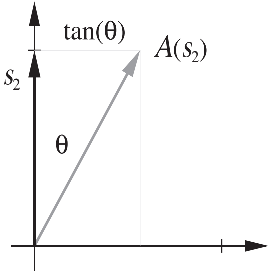
If we draw a line on the bottom card and then skew the deck, the line will not change.



$$A(s_1) = 1s_1 + 0s_2 = \sum_{i=1}^{2} a_{i1}s_i = a_{11}s_1 + a_{21}s_2$$

This gives us the first column of the matrix.

Next, we skew s_2:



$$A(s_2) = \tan(\theta)s_1 + 1s_2 = \sum_{i=1}^{2} a_{i2}s_i = a_{12}s_1 + a_{22}s_2$$

This gives us the second column of the matrix.

The matrix of the transformation is:

$$\mathbf{A} = \begin{bmatrix} 1 & \tan(\theta) \\ 0 & 1 \end{bmatrix}$$

Consider the linear transformation $\mathcal{A}: X \rightarrow Y$. Let $\{v_1, v_2, \ldots, v_n\}$ be a basis for X, and let $\{u_1, u_2, \ldots, u_m\}$ be a basis for Y.

$$x = \sum_{i=1}^{n} x_i v_i \qquad\qquad y = \sum_{i=1}^{m} y_i u_i$$

$$\mathcal{A}(x) = y$$

The matrix representation is:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

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

Now let's consider different basis sets. Let $\{t_1, t_2, ..., t_n\}$ be a basis for X, and let $\{w_1, w_2, ..., w_m\}$ be a basis for Y.

$$x = \sum_{i=1}^{n} x'_i t_i \qquad y = \sum_{i=1}^{m} y'_i w_i$$

The new matrix representation is:

$$\begin{bmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ a'_{21} & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & & \vdots \\ a'_{m1} & a'_{m2} & \cdots & a'_{mn} \end{bmatrix} \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix} = \begin{bmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_m \end{bmatrix}$$

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

Expand t_i in terms of the original basis vectors for X.

$$t_i = \sum_{j=1}^{n} t_{ji} v_j \qquad \mathbf{t}_i = \begin{bmatrix} t_{1i} \\ t_{2i} \\ \vdots \\ t_{ni} \end{bmatrix}$$

Expand w_i in terms of the original basis vectors for Y.

$$w_i = \sum_{j=1}^{m} w_{ji} u_j \qquad \mathbf{w}_i = \begin{bmatrix} w_{1i} \\ w_{2i} \\ \vdots \\ w_{mi} \end{bmatrix}$$

How A and A' related?

$$\mathbf{B}_t = \begin{bmatrix} \mathbf{t}_1 & \mathbf{t}_2 & \dots & \mathbf{t}_n \end{bmatrix} \qquad \mathbf{x} = x'_1\mathbf{t}_1 + x'_2\mathbf{t}_2 + \cdots + x'_n\mathbf{t}_n = \mathbf{B}_t\mathbf{x}'$$

$$\mathbf{B}_w = \begin{bmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \dots & \mathbf{w}_m \end{bmatrix} \qquad \mathbf{y} = \mathbf{B}_w\mathbf{y}'$$

$$\mathbf{A}\mathbf{x} = \mathbf{y} \quad \Rightarrow \quad \mathbf{A}\mathbf{B}_t\mathbf{x}' = \mathbf{B}_w\mathbf{y}'$$

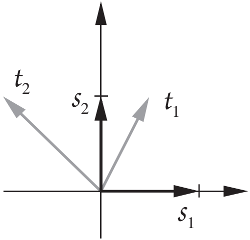
$$[\mathbf{B}_w^{-1}\mathbf{A}\mathbf{B}_t]\mathbf{x}' = \mathbf{y}'$$

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

$$\Rightarrow \quad \boxed{\mathbf{A}' = [\mathbf{B}_w^{-1}\mathbf{A}\mathbf{B}_t]}$$

Similarity Transform

Take the skewing problem described previously, and find the new matrix representation using the basis set $\{s_1, s_2\}$.



$$t_1 = 0.5 s_1 + s_2$$

$$t_2 = -s_1 + s_2$$

$$\mathbf{t}_1 = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} \qquad \mathbf{t}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$\Rightarrow \quad \mathbf{B}_t = \begin{bmatrix} \mathbf{t}_1 & \mathbf{t}_2 \end{bmatrix} = \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix} \qquad \mathbf{B}_w = \mathbf{B}_t = \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix}$$

(Same basis for domain and range.)

Example (2)

$$\mathbf{A}' = [\mathbf{B}_w^{-1}\mathbf{A}\mathbf{B}_t] = \begin{bmatrix} 2/3 & 2/3 \\ -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1 & \tan\theta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix}$$

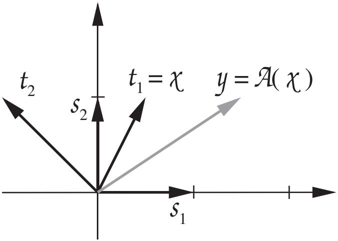
$$\mathbf{A}' = \begin{bmatrix} (2/3)\tan\theta + 1 & (2/3)\tan\theta \\ (-2/3)\tan\theta & (-2/3)\tan\theta + 1 \end{bmatrix}$$

For $\theta = 45°$:

$$\mathbf{A}' = \begin{bmatrix} 5/3 & 2/3 \\ -2/3 & 1/3 \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Try a test vector: $\mathbf{x} = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}$ $\mathbf{x}' = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$$\mathbf{y} = \mathbf{A}\mathbf{x} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix} \qquad \mathbf{y}' = \mathbf{A}'\mathbf{x}' = \begin{bmatrix} 5/3 & 2/3 \\ -2/3 & 1/3 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5/3 \\ -2/3 \end{bmatrix}$$



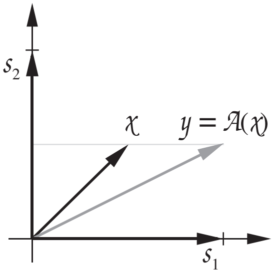
Check using reciprocal basis vectors:

$$\mathbf{y}' = \mathbf{B}^{-1}\mathbf{y} = \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix}^{-1}\begin{bmatrix} 1.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/3 & 2/3 \\ -2/3 & 1/3 \end{bmatrix}\begin{bmatrix} 1.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 5/3 \\ -2/3 \end{bmatrix}$$

Let $\mathcal{A}: X \to X$ be a linear transformation. Those vectors $z \in X$, which are not equal to zero, and those scalars λ which satisfy

$$\mathcal{A}(z) = \lambda z$$

are called eigenvectors and eigenvalues, respectively.



Can you find an eigenvector for this transformation?

$$\mathbf{A}\mathbf{z} = \lambda\mathbf{z}$$

$$[\mathbf{A} - \lambda\mathbf{I}]\mathbf{z} = \mathbf{0} \quad \Rightarrow \quad |[\mathbf{A} - \lambda\mathbf{I}]| = 0$$

Skewing example (45°):

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad \left|\begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix}\right| = 0 \qquad (1-\lambda)^2 = 0 \qquad \begin{matrix} \lambda_1 = 1 \\ \lambda_2 = 1 \end{matrix}$$

$$\begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix}\mathbf{z} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\mathbf{z}_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} z_{11} \\ z_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad z_{21} = 0 \qquad \mathbf{z}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

For this transformation there is only one eigenvector.

Perform a change of basis (similarity transformation) using the eigenvectors as the basis vectors. If the eigenvalues are distinct, the new matrix will be diagonal.

$$\mathbf{B} = \begin{bmatrix} \mathbf{z}_1 & \mathbf{z}_2 & \dots & \mathbf{z}_n \end{bmatrix}$$

$\{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n\}$ Eigenvectors

$\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ Eigenvalues

$$[\mathbf{B}^{-1}\mathbf{A}\mathbf{B}] = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

Example

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

$$\left| \begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} \right| = 0 \qquad \lambda^2 - 2\lambda = (\lambda)(\lambda-2) = 0 \qquad \begin{matrix} \lambda_1 = 0 \\ \lambda_2 = 2 \end{matrix} \qquad \begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} \mathbf{z} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\lambda_1 = 0 \quad \Rightarrow \quad \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \mathbf{z}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} z_{11} \\ z_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad z_{21} = -z_{11} \qquad \mathbf{z}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\lambda_2 = 2 \quad \Rightarrow \quad \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \mathbf{z}_1 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} z_{12} \\ z_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad z_{22} = z_{12} \qquad \mathbf{z}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Diagonal Form: $\mathbf{A}' = [\mathbf{B}^{-1}\mathbf{A}\mathbf{B}] = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$