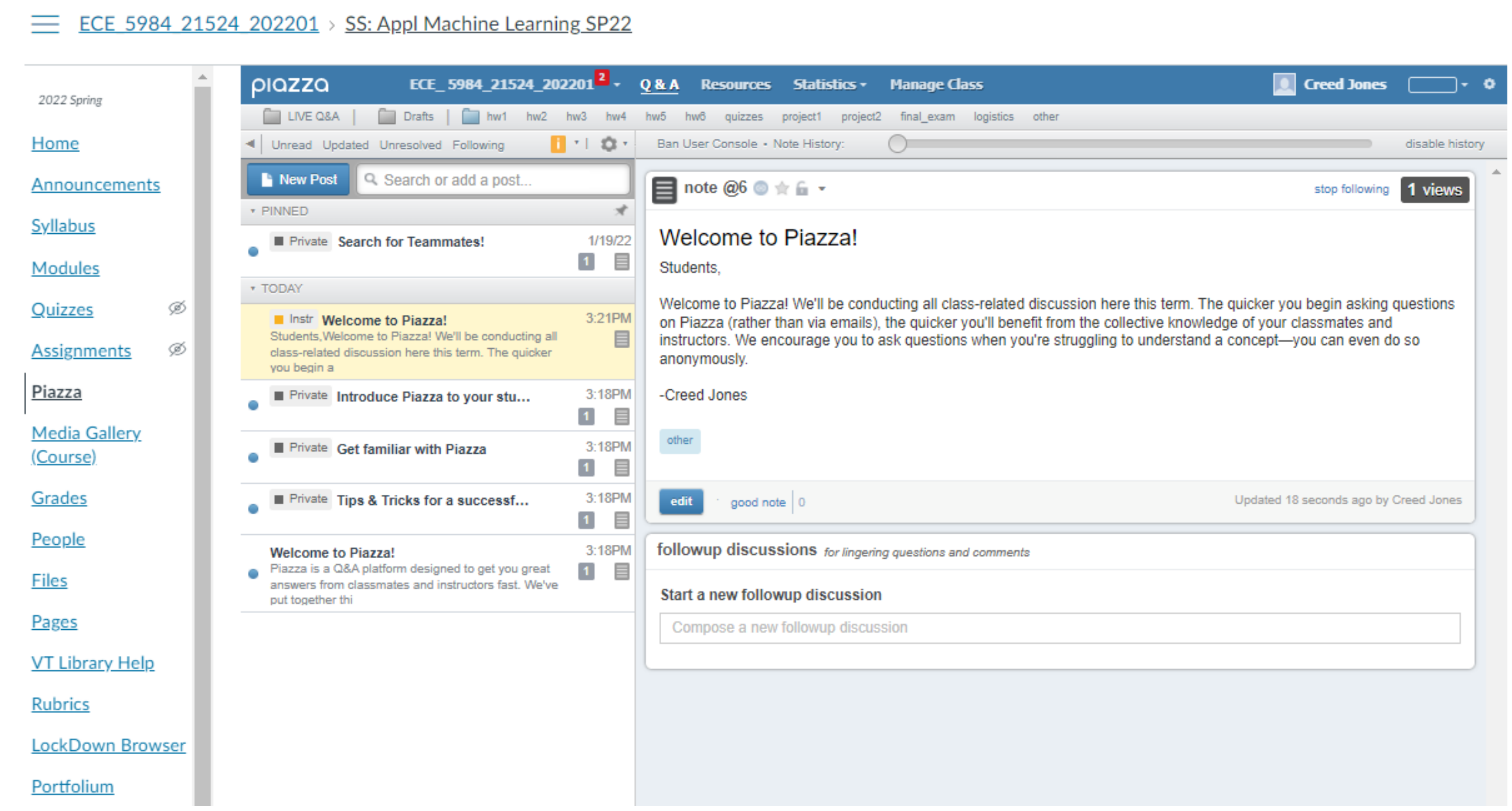
ECE5984 – Applications of Machine Learning
Lecture 2 – Review of Linear Algebra

Creed Jones, PhD

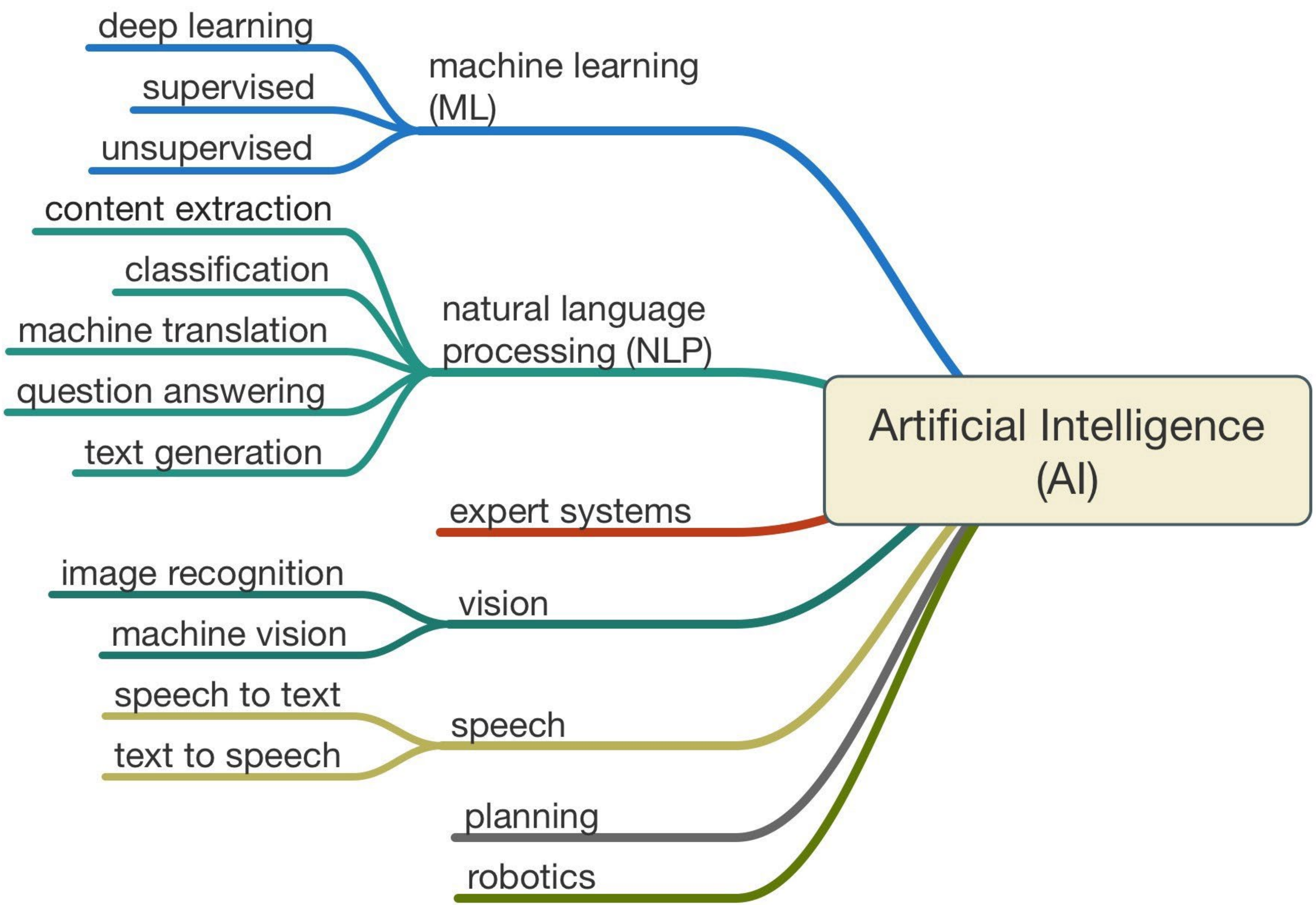
Course Update

- Graduate Teaching Assistant – Ashley Smith
  - Office hours: Tuesday and Thursday, 10 AM to noon
- Quiz 1 will be on Thursday, January 27
  - On lectures 1-3
  - Must be taken between 12 noon and 6 PM
  - 20 minute time limit
- Homework 1 will be posted early next week
  - Due on Tuesday, February 8
- If you send me an email on the course, please put “ECE5984” in the subject line!

I use Piazza for questions on assignments, lectures and so on



- Accessible through the menu in Canvas
- It's a great place to ask and to answer questions
- I allow anonymous posting
- I do not allow private posting



From Mira Bhattacharya –
https://medium.com/@mirabhattacharya/the-biological-taxonomy-of-artificial-intelligence-a-crash-course-3cc1a2d7ee3c

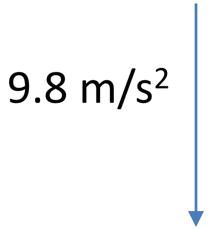
Today's Objectives

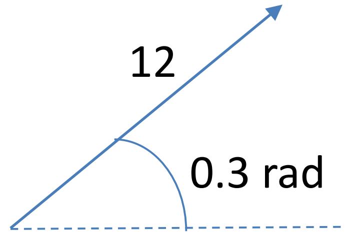
- Vectors
  - Vector products
  - Vector differentiation
  - Vector norms
- Matrices
  - Matrix products
  - Inversion and singularity
  - Special matrices
  - Solving systems of equations
  - Matrix decomposition
  - Eigendecomposition

VECTORS

A vector is a mathematical quantity with both <u>magnitude</u> and <u>direction</u>

- In early physics courses, we represent them as arrows





- More usefully, we can represent them as a collection of scalars that are the *components* of the vector in multiple *dimensions*

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \begin{bmatrix} 2 & -3 & 0.71 \end{bmatrix}$$

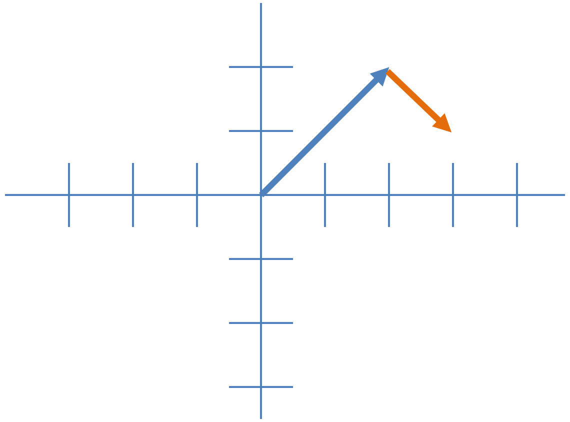
A vector also obeys the governing rules of *vector addition*

- Addition of two vectors *a* and *b* results in a vector (call it *c*)

- *c* can be found by addition of the components:

$$a = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \; b = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, c = a + b = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

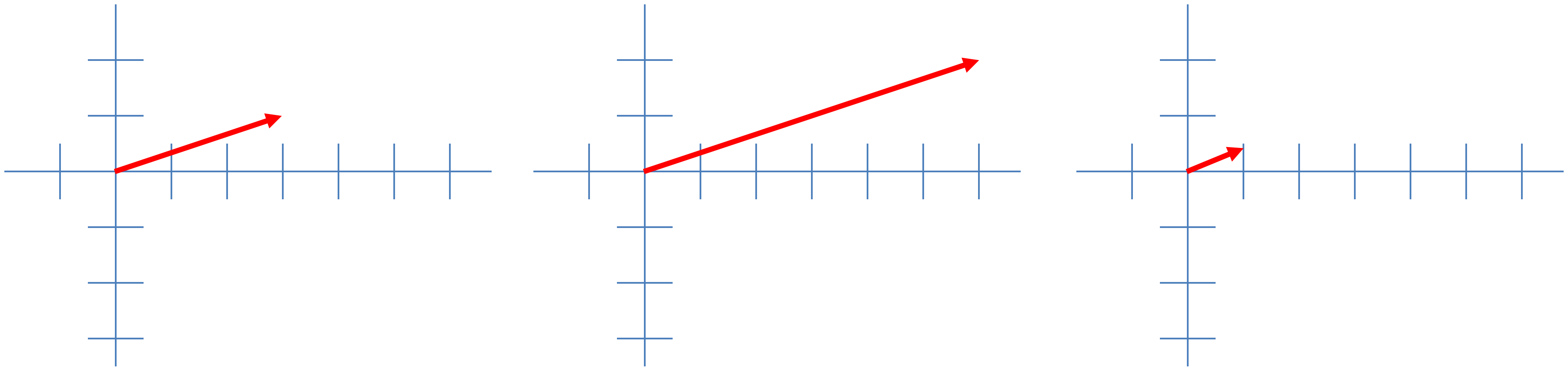
- *c* can also be found by "head-to-tail" addition:



Multiplying a vector by a scalar is defined as the product of that scalar with each component – the magnitude changes <u>but not the direction</u>

- Scalar multiplication of a vector

$$a = \begin{bmatrix} 3 \\ 1 \end{bmatrix},\ b = 2a = \begin{bmatrix} 2(3) \\ 2(1) \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix},\ c = \frac{1}{3}a = \begin{bmatrix} \frac{1}{3}(3) \\ \frac{1}{3}(1) \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{1}{3} \end{bmatrix}$$

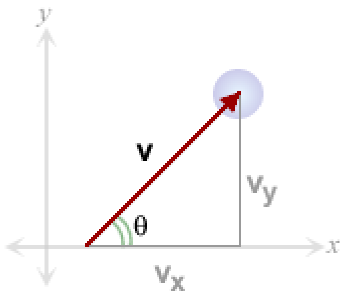


Given a coordinate system, the relationships between components and angles are simple trigonometry

- In 2D,

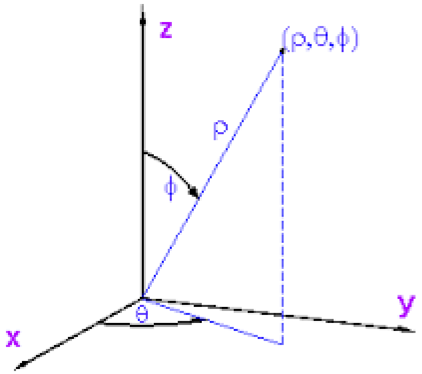
$$v_x = V\cos\theta$$
$$v_y = V\sin\theta$$

signs are important



$$\begin{bmatrix} V\cos\theta \\ V\sin\theta \end{bmatrix}$$

- In 3D,



$$\begin{bmatrix} V\cos\theta\sin\phi \\ V\sin\theta\sin\phi \\ V\cos\phi \end{bmatrix}$$

Scalar (dot) product and Vector (cross) product

Dot Product

- $A \cdot B = ABcos\theta$, where θ is the angle between A and B
- $A \cdot B = [a_1 \quad a_2 \quad a_3] \cdot \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = a_1b_1 + a_2b_2 + a_3b_3$

$$\boldsymbol{A \cdot B = B \cdot A}$$

$$\boldsymbol{A \times B = -B \times A}$$

Cross Product

- $\|A \times B\| = ABsin\theta$, $\measuredangle(A \times B)$ is perpendicular to both A and B (obeys the *right hand rule)*
- $A \times B = [a_1 \quad a_2 \quad a_3] \times \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{vmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$,

The lines indicate the *determinant* of the matrix, i, j, and k are unit vectors in the directions of the x, y and z axes

The typical vector cross product is only defined for vectors of three dimensions!

- 3D: $A \times B = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{vmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$

- 4D: $A \times B = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{vmatrix} i & j & k & l \\ a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \end{vmatrix} = ??????$

- The cross product has been generalized in a number of ways:
  - Quaternions and octonions
  - Lie algebra
  - External product
  - …

We will soon see (or recall from earlier courses) the process of finding the *determinant* of a matrix

$$|A| = det(A) = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

$$= a(ei - fh) - b(di - fg) + c(dh - eg)$$

- Larger matrices can be handled using the same reduction procedure

Let’s try a few questions

Let $A = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $C = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$

1. What’s $A \cdot B$?

2. What’s $A \times B$?

3. What’s $(A \times B) \cdot C$?

4. What’s $(A \times B) \times C$?

Let’s try a few questions

Let $A = \begin{bmatrix}1\\2\\3\end{bmatrix}$, $B = \begin{bmatrix}1\\0\\-1\end{bmatrix}$, $C = \begin{bmatrix}2\\1\\0\end{bmatrix}$

1. What’s $A \cdot B$? $A \cdot B = \begin{bmatrix}1 & 2 & 3\end{bmatrix} \cdot \begin{bmatrix}1\\0\\-1\end{bmatrix} = 1 \cdot 1 + 2 \cdot 0 + 3 \cdot (-1) = -2$

2. What’s $A \times B$? $A \times B = \begin{vmatrix} i & j & k \\ 1 & 2 & 3 \\ 1 & 0 & -1 \end{vmatrix} = -2i - (-4)j - 2k = \begin{bmatrix}-2\\4\\-2\end{bmatrix}$

3. What’s $(A \times B) \cdot C$? $= \begin{bmatrix}-2 & 4 & -2\end{bmatrix} \cdot \begin{bmatrix}2\\1\\0\end{bmatrix} = -4 + 4 + 0 = 0$

4. What’s $(A \times B) \times C$? $= \begin{vmatrix} i & j & k \\ -2 & 4 & -2 \\ 2 & 1 & 0 \end{vmatrix} = 2i - (4)j - 10k = \begin{bmatrix}2\\-4\\-10\end{bmatrix}$

Differentiation is defined on vectors, as the differentiation of the components

$$\frac{d}{dt}A(t) = i\frac{d}{dt}a_x(t) + j\frac{d}{dt}a_y(t) + k\frac{d}{dt}a_z(t) = \begin{bmatrix} \frac{da_x}{dt} \\ \frac{da_y}{dt} \\ \frac{da_z}{dt} \end{bmatrix} = \frac{d}{dt}\begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}$$

- In general, the derivative of a vector does not point in the same direction as the vector itself
  - Just as the derivative of a scalar function doesn't necessarily have the same value as the function

Properties for vector differentiation

- If $\boldsymbol{A}$ and $\boldsymbol{B}$ are vectors that are functions of t,
- $\frac{d}{dt}\{c\boldsymbol{A}\} = c\frac{d\boldsymbol{A}}{dt}$
- $\frac{d}{dt}\{\boldsymbol{A}+\boldsymbol{B}\} = \frac{d\boldsymbol{A}}{dt} + \frac{d\boldsymbol{B}}{dt}$
- $\frac{d}{dt}\{\boldsymbol{A}\cdot\boldsymbol{B}\} = \boldsymbol{A}\cdot\frac{d\boldsymbol{B}}{dt} + \frac{d\boldsymbol{A}}{dt}\cdot\boldsymbol{B}$
- $\frac{d}{dt}\{\boldsymbol{A}\times\boldsymbol{B}\} = \boldsymbol{A}\times\frac{d\boldsymbol{B}}{dt} + \frac{d\boldsymbol{A}}{dt}\times\boldsymbol{B}$
- If $\boldsymbol{A}(t)\cdot\boldsymbol{A}(t) = constant$, then $\boldsymbol{A}$ and $\frac{d\boldsymbol{A}}{dt}$ are perpendicular – that is, $\boldsymbol{A}(t)\cdot\frac{d}{dt}\boldsymbol{A}(t) = 0$

Partial differentiation of vectors works at the component level

Let $\boldsymbol{A}$ be a function of u and v:

- $$\frac{\partial \boldsymbol{A}}{\partial u} = i\frac{\partial A_x}{\partial u} + j\frac{\partial A_y}{\partial u} + k\frac{\partial A_z}{\partial u} = \begin{bmatrix} \frac{\partial A_x}{\partial u} \\ \frac{\partial A_y}{\partial u} \\ \frac{\partial A_z}{\partial u} \end{bmatrix}$$

- $$\frac{\partial^2 \boldsymbol{A}}{\partial u \partial v} = i\frac{\partial^2 A_x}{\partial u \partial v} + j\frac{\partial^2 A_y}{\partial u \partial v} + k\frac{\partial^2 A_z}{\partial u \partial v} = \begin{bmatrix} \frac{\partial^2 A_x}{\partial u \partial v} \\ \frac{\partial^2 A_y}{\partial u \partial v} \\ \frac{\partial^2 A_z}{\partial u \partial v} \end{bmatrix},$$ and so on…

Integration of a vector function over an independent variable can be done by straightforward integration of the components

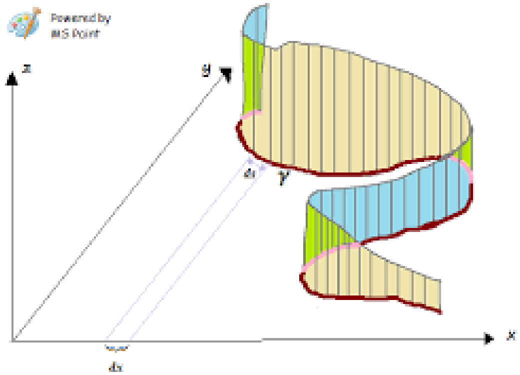
- For a vector function of time:

$$\vec{f}(t) = f_x(t)\vec{\imath} + f_y(t)\vec{\jmath} + f_z(t)\vec{k}$$

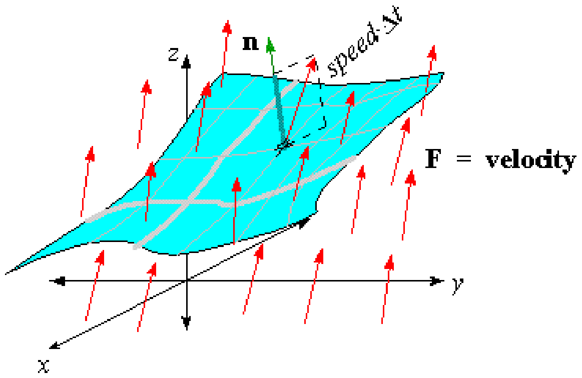
the time-based integral is:

$$\int \vec{f}(t)dt = \int f_x(t)dt\,\vec{\imath} + \int f_y(t)dt\,\vec{\jmath} + \int f_z(t)dt\,\vec{k}$$

Integration of a vector function over its component variables is more complex; there are two common types: line and surface integrals



- A line integral (poorly named) is the integration of a vector function along some path in the coordinate system of the vector



- A surface integral is the integration of a vector function across an entire surface in the vector space

We can define important scalar measures on a vector – commonly called *norms*

A norm is a function that assigns a scalar to a vector
Another common notation for a norm on the vector $\vec{v}$ is $\|\vec{v}\|$

There are some important properties on the norm n for a vector $\vec{v}$

$n(\vec{v}) = 0 \ \ iff \ \ \vec{v} = \vec{0}$	$\|\vec{v}\| = 0 \ \ iff \ \ \vec{v} = \vec{0}$	: only the zero vector has a zero norm
$n(\vec{v}) > 0 \ \ iff \ \ \vec{v} \neq \vec{0}$	$\|\vec{v}\| > 0 \ \ iff \ \ \vec{v} \neq \vec{0}$	: a norm is nonnegative
$n(a\vec{v}) = an(\vec{v})$	$\|a\vec{v}\| = a\|\vec{v}\|$	: norms scale linearly
$n(\vec{v} + \vec{w}) \leq n(\vec{v}) + n(\vec{w})$	$\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$	: triangle inequality

Euclidean Distance, also called the L^2 norm, is a valid norm

- In 2 dimensions: $L^2 = \sqrt{v_x^2 + v_y^2}$
- In 3 dimensions: $L^2 = \sqrt{v_x^2 + v_y^2 + v_z^2}$
- Generally: $L^2 = \sqrt{\sum_{i=1}^{k} v_i^2}$

The Euclidean norm is a valid norm:

- $n(\vec{v}) = 0 \;\; iff \;\; \vec{v} = \vec{0}$: $\sqrt{\sum_{i=1}^{k} v_i^2} = 0 \; iff \; v_i = 0 \; \forall \, i$
- $n(\vec{v}) > 0 \;\; iff \;\; \vec{v} \neq \vec{0}$: $\sqrt{\sum_{i=1}^{k} v_i^2} \geq 0 \; \forall \, v_i$
- $n(a\vec{v}) = an(\vec{v})$: $\sqrt{\sum_{i=1}^{k} a^2 v_i^2} = \sqrt{a^2 \sum_{i=1}^{k} v_i^2} = a\sqrt{\sum_{i=1}^{k} v_i^2}$
- $n(\vec{v} + \vec{w}) \leq n(\vec{v}) + n(\vec{w})$: can be shown by simple algebra

There are several useful vector norms in use

- $L^2 = \sqrt[2]{\sum_{i=1}^{k} v_i^2}$
- $L^3 = \sqrt[3]{\sum_{i=1}^{k} v_i^3}$
- $L^\infty = \sqrt[\infty]{\sum_{i=1}^{k} v_i^\infty}$
- Manhattan (city block) norm: $\sum_{i=1}^{k} |v_i|$

MATRICES

A Matrix is a rectangular array of numeric expressions, generally describing interrelated systems or quantities

- A matrix has an *order* – “n by m” means n rows and m columns
  - This matrix is 2 by 3: $\begin{bmatrix} 7 & x & 2.3 \\ 0 & 1+7j & \sin(x) \end{bmatrix}$
- Conventionally, variables representing matrices are represented in caps: $\boldsymbol{A} = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$
- Individual terms are shown using subscripts, row first:
- $a_{11} = 1, \;\; a_{21} = 2, \;\; a_{12} = 4, \;\; a_{22} = 3$

The determinant of a matrix is a scalar calculated from all elements of the matrix

- The determinant of a 2x2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$
- The determinant of a 3x3 matrix is the sum of the determinants of three *minors* of the matrix, times alternating +1 and -1 coefficients:
- $det\left(\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}\right)$

$$= (+1)a \cdot det\left(\begin{bmatrix} e & f \\ h & i \end{bmatrix}\right) + (-1)b \cdot det\left(\begin{bmatrix} d & f \\ g & i \end{bmatrix}\right) + (+1)c \cdot det\left(\begin{bmatrix} d & e \\ g & h \end{bmatrix}\right)$$

$$= (+1)a(ei - fh) + (-1)b(di - fg) + (+1)c(dh - eg)$$

Some properties of determinants

1. $\det I = |I| = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1$
2. If you exchange two rows of a matrix, you reverse the sign of its determinant.
3. If we multiply one row of a matrix by t, the determinant is multiplied by t
4. If two rows of a matrix are equal, its determinant is zero.
5. If $i \neq j$, subtracting t times row i from row j doesn't change the determinant.

The result of multiplying two matrices is composed of the sum of the product of row terms and column terms

$$\boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{C}$$

$$\boldsymbol{c_{xy}} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{r}} \boldsymbol{a_{xi} b_{iy}}$$

$$\boldsymbol{Note}:$$

$$\boldsymbol{A} \cdot \boldsymbol{B} \neq \boldsymbol{B} \cdot \boldsymbol{A}$$

The inner dimension of matrices to be multiplied must match

If $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$, and $B = \begin{bmatrix} 4 & 5 \\ 6 & 7 \end{bmatrix}$,

$$\boldsymbol{A} \cdot \boldsymbol{B} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 4 & 5 \\ 6 & 7 \end{bmatrix} = \begin{bmatrix} (1 \cdot 4 + 2 \cdot 6) & (1 \cdot 5 + 2 \cdot 7) \\ (0 \cdot 4 + 3 \cdot 6) & (0 \cdot 5 + 3 \cdot 7) \end{bmatrix} = \begin{bmatrix} 16 & 19 \\ 18 & 21 \end{bmatrix}$$

When multiplying matrices, the *inner dimension* must be the same or multiplication is not defined

$$\boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{C}\text{;}\ \boldsymbol{c_{xy}} = \sum_{i=1}^{r} \boldsymbol{a_{xi} b_{iy}}$$

If $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 0 \end{bmatrix}$, and $B = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 3 & 2 & 1 \\ 0 & 1 & 2 & 0 \end{bmatrix}$,

$$\boldsymbol{A} \cdot \boldsymbol{B} =$$

$$\begin{bmatrix} (1 \cdot 1 + 2 \cdot 1 + 1 \cdot 0) & (1 \cdot 1 + 2 \cdot 3 + 1 \cdot 1) & (1 \cdot 1 + 2 \cdot 2 + 1 \cdot 2) & (1 \cdot 0 + 2 \cdot 1 + 1 \cdot 0) \\ (1 \cdot 1 + 3 \cdot 1 + 0 \cdot 0) & (1 \cdot 1 + 3 \cdot 3 + 0 \cdot 0) & (1 \cdot 1 + 3 \cdot 2 + 0 \cdot 2) & (1 \cdot 0 + 3 \cdot 1 + 0 \cdot 0) \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 8 & 7 & 2 \\ 4 & 10 & 7 & 3 \end{bmatrix}$$

$$\boldsymbol{B} \cdot \boldsymbol{A} = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 3 & 2 & 1 \\ 0 & 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 0 \end{bmatrix} = ???$$

The Inverse of a matrix $\boldsymbol{A}$ is the solution to the equation $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix

- $\boldsymbol{I}$ is the square matrix with 1s in diagonal spots and 0s in the non-diagonal entries
  - A 3x3 identity matrix is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
  - Clearly, identity matrices must be square

- Inversions of matrices:
- $\begin{bmatrix} 1 & 3 & 5 \\ 2 & 5 & 3 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} {}^{1}/_{3} & {}^{-7}/_{3} & {}^{16}/_{3} \\ {}^{-1}/_{3} & {}^{4}/_{3} & {}^{-7}/_{3} \\ {}^{1}/_{3} & {}^{-1}/_{3} & {}^{1}/_{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Let's do a few examples:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 3 \end{bmatrix}; \; B = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix}$$

- $3B = 3$
- $AB =$
- $BA =$
- $\det A =$
- $\det B =$

Let's do a few examples:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 3 \end{bmatrix}; \quad B = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix}$$

- $3B = 3\begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix} = \begin{bmatrix} -3 & 6 \\ 6 & -12 \end{bmatrix}$
- $AB = \begin{bmatrix} 1 & 3 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix} = \begin{bmatrix} 1(-1)+3(2) & 1(2)+3(-4) \\ 2(-1)+3(2) & 2(2)+3(-4) \end{bmatrix} = \begin{bmatrix} 5 & -10 \\ 4 & -8 \end{bmatrix}$
- $BA = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} -1(1)+2(2) & -1(3)+2(3) \\ 2(1)+-4(2) & 2(3)+-4(3) \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ -6 & -6 \end{bmatrix}$

- $\det A = \begin{vmatrix} 1 & 3 \\ 2 & 3 \end{vmatrix} = 1(3) - 3(2) = -3$
- $\det B = \begin{vmatrix} -1 & 2 \\ 2 & -4 \end{vmatrix} = -1(-4) - 2(2) = 0$

A Singular Matrix is a square matrix that does not have an inverse

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 4 \\ 1 & 2 & 1 \end{bmatrix}$$, for example

Singular matrices always have zero determinants (and matrices with zero determinants are always singular):

$$\det\left(\begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 4 \\ 1 & 2 & 1 \end{bmatrix}\right) = 1(3\cdot 1 - 4\cdot 2) - 2(2\cdot 1 - 4\cdot 1) + 1(2\cdot 2 - 3\cdot 1)$$

$$= 3 - 8 - 4 + 8 - 4 - 3 = 0$$

The Rank of a Matrix is the number of independent rows or columns it contains

- Formally, the rank is the dimension of the largest non-singular submatrix
- The rank of a matrix may be less than its dimension
  - If so, it's called *rank-deficient*
  - If the rank is the dimension, the matrix is called *full-rank*

- $rk\begin{pmatrix} 1 & 4 & 2 \\ 6 & 0 & 1 \\ 3 & 2 & 3 \end{pmatrix} = 3$
- $rk\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 3 & 2 & 3 \end{pmatrix} = 2$, because row 2 is 2 times row 1

Elementary row operations on a matrix result in a *row-equivalent* matrix

Elementary row operations include:

- Interchanging two rows
- Multiplying each element of a row by a (non-zero) constant
- Adding or subtracting one row from another, element by element

The corresponding *elementary column operations* result in a *column-equivalent matrix*

Any matrix has the same rank as its row- and column-equivalent matrices

Row equivalence and column equivalence (denoted by ~) are equivalence relations

They are:

- Reflexive: A~A (by definition)
- Symmetric: if A~B, then B~A (because row operations are reversible)
- Transitive: if A~B and B~C, then A~C

Simultaneous linear equations in several variables can be nicely expressed in a matrix equation

- $-x - 2y + 3z = 4$
- $2x + y + z = 0$
- $3x - 4y - 2z = -1$

$$\begin{bmatrix} -1 & -2 & 3 \\ 2 & 1 & 1 \\ 3 & -4 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ -1 \end{bmatrix}$$

- If a solution exists, the equations are said to be *consistent*

- Conversely, if we can test a matrix for consistency of the equations, we can determine if a solution exists

Solutions of Equations – Inverse Method

Consider a set of equations and the matrix form:

- $3x + 2y - z = 4$
- $2x - y + 2z = 10$
- $x \ - 3y - 4z = 5$

$$A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 & 2 & -1 \\ 2 & -1 & 2 \\ 1 & -3 & -4 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 10 \\ 5 \end{bmatrix}$$

- $A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = B$ $\qquad A^{-1}A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1}B$

- $A^{-1} = \frac{1}{55}\begin{bmatrix} 10 & 11 & 3 \\ 10 & -11 & -8 \\ -5 & 11 & -7 \end{bmatrix}$

Application of the Inverse Method

Consider a set of equations and the matrix form:

- $3x + 2y - z = 4$
- $2x - y + 2z = 10$
- $x \ - 3y - 4z = 5$

$$A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 & 2 & -1 \\ 2 & -1 & 2 \\ 1 & -3 & -4 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 10 \\ 5 \end{bmatrix}$$

- $A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = B \qquad A^{-1}A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1}B$

- $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1}B = \frac{1}{55}\begin{bmatrix} 10 & 11 & 3 \\ 10 & -11 & -8 \\ -5 & 11 & -7 \end{bmatrix}\begin{bmatrix} 4 \\ 10 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$

Here are a couple of nice online resources for manipulations of matrices

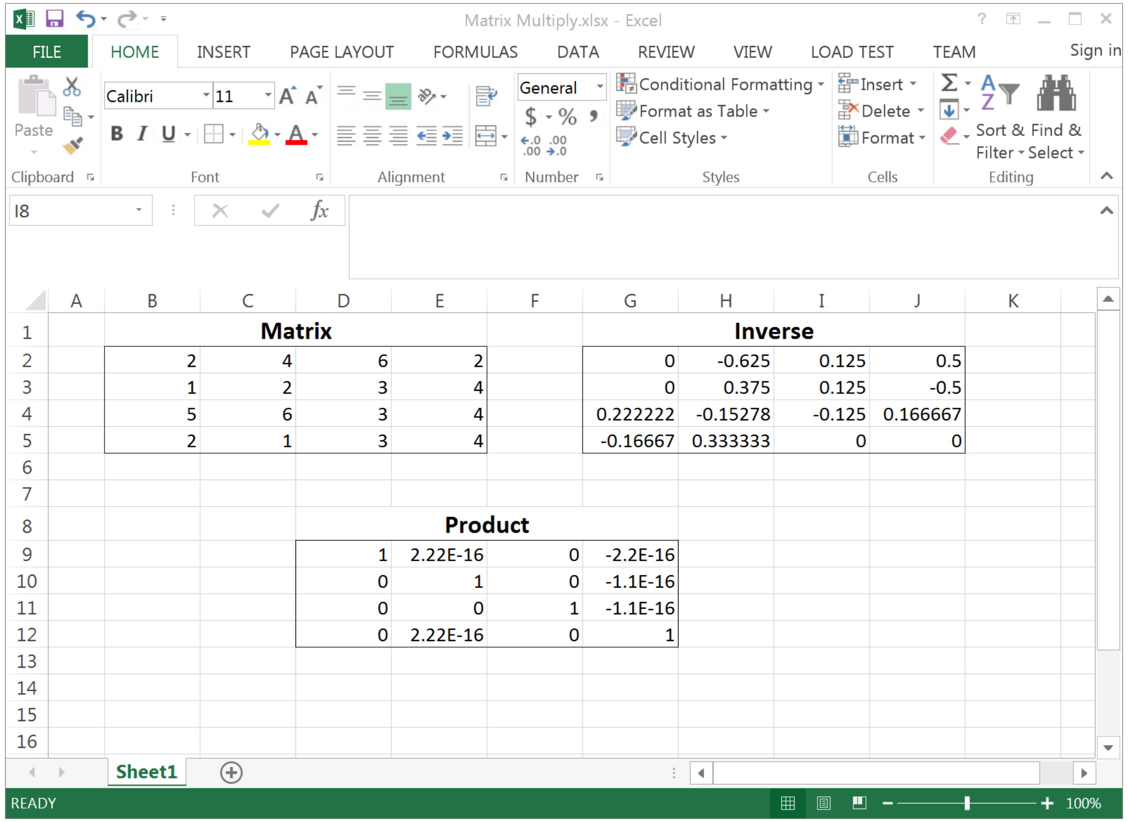
- The Online Matrix Calculator (https://www.calculator.net/matrix-calculator.html) performs common operations on matrices up to 10 by 10

- Dr. Jim Carrell, emeritus professor at the University of British Columbia has written a good book on Linear Algebra and Matrices: it's available at http://www.math.ubc.ca/~carrell/NB.pdf

Microsoft Excel provides some nice facilities for simple matrix operations

MATLAB is the standard for manipulating matrices, but Excel can do a few things…

In Python, we will see that Numpy also provides linear algebra support



Matrix

2	4	6	2
1	2	3	4
5	6	3	4
2	1	3	4

Inverse

0	-0.625	0.125	0.5
0	0.375	0.125	-0.5
0.222222	-0.15278	-0.125	0.166667
-0.16667	0.333333	0	0

Product

1	2.22E-16	0	-2.2E-16
0	1	0	-1.1E-16
0	0	1	-1.1E-16
0	2.22E-16	0	1

The *transpose* of a matrix is formed by interchanging elements across the main diagonal

- The transpose of $\boldsymbol{A}$ is written as $\boldsymbol{A}^{\boldsymbol{T}}$
- If $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$, then $\boldsymbol{A}^{\boldsymbol{T}} = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$

$$a_{ij}^{T} = a_{ji}$$

The *conjugate transpose* of a matrix is formed by interchanging elements across the main diagonal, then forming the complex conjugate of each

- The conjugate transpose of $\boldsymbol{A}$ is written as $\boldsymbol{A}^*$
  - sometimes it's written A^H
- If $\boldsymbol{A} = \begin{bmatrix} 1-j & 2+2j & 3 \\ 4 & 1 & 3 \\ 0 & 8-3j & 9 \end{bmatrix}$, then $\boldsymbol{A}^* = \begin{bmatrix} 1+j & 4 & 0 \\ 2-2j & 1 & 8+3j \\ 3 & 3 & 9 \end{bmatrix}$

$$a_{ij}^* = \left(a_{ji}\right)^*$$

where the * indicates complex conjugation (negating the imaginary parts but not the real parts)

$$(a+bj)^* = a - bj$$

A number of important types of matrices have been defined; they're useful in advanced operations on matrices

- Symmetric matrix: a <u>real</u> matrix $\boldsymbol{A}$ is symmetric if $\boldsymbol{A} = \boldsymbol{A}^{\boldsymbol{T}}$

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & 7 \\ 1 & 7 & 4 \end{bmatrix}$$

- Hermitian matrix: a <u>complex</u> matrix $\boldsymbol{A}$ is Hermitian if $\boldsymbol{A} = \boldsymbol{A}^{*}$

$$\begin{bmatrix} 1 & 2+\mathrm{j} & 1 \\ 2-j & 5 & -3\mathrm{j} \\ 1 & 3\mathrm{j} & 4 \end{bmatrix}$$

- Orthogonal matrix: a real matrix $\boldsymbol{A}$ is orthogonal if $\boldsymbol{A}^{-\boldsymbol{1}} = \boldsymbol{A}^{\boldsymbol{T}}$

$$\begin{bmatrix} cos\theta & sin\theta \\ sin\theta & -cos\theta \end{bmatrix} \begin{bmatrix} cos\theta & sin\theta \\ sin\theta & -cos\theta \end{bmatrix} =$$

$$\begin{bmatrix} cos^2\theta + sin^2\theta & cos\theta sin\theta - sin\theta cos\theta \\ sin\theta cos\theta - cos\theta sin\theta & sin^2\theta + cos^2\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Orthogonal Matrices are especially important -

The columns and rows are all orthogonal (linearly independent) unit vectors; therefore:

- $\boldsymbol{A}$ is invertible; there is an $\boldsymbol{A}^{-1}$
- If $\boldsymbol{A}^{-1} = \boldsymbol{A}^{T}$, then $\boldsymbol{A}^{T}\boldsymbol{A} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$
- The determinant of an orthogonal matrix is either 1 or -1
- The columns and rows are orthogonal unit vectors

A positive definite matrix gives nonnegative results when pre- and post-multiplied by any vector in n-space

- Positive definite matrix: a matrix $\boldsymbol{A}$ is positive definite if $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{0}$ for <u>any</u> real-valued vector x

$$\boldsymbol{A} \text{ is positive definite iff } \forall x \in R^n, x^T\boldsymbol{A}x \geq 0$$

- Take $\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$, pre- and post-multiply an arbitrary vector:

$$[x \quad y \quad z]\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = [x \quad y \quad z]\begin{bmatrix} 2x - y \\ -x + 2y - z \\ -y + 2z \end{bmatrix}$$

$$= 2x^2 - xy - xy + 2y^2 - yz - yz + 2z^2 = 2x^2 - 2xy + 2y^2 - 2yz + 2z^2$$

$$= x^2 + (x^2 - 2x + y^2) + (y^2 - 2yz + z^2) + z^2$$

$$= x^2 + (x - y)^2 + (y - z)^2 + z^2$$

A positive definite matrix gives nonnegative results when pre- and post-multiplied by any vector in n-space

- Positive definite matrix: a matrix $\boldsymbol{A}$ is positive definite if $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{0}$ for <u>any</u> real-valued vector x

$$\boldsymbol{A} \text{ is positive definite iff } \forall x \in R^n, x^T\boldsymbol{A}x \geq 0$$

- Take $\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$, pre- and post-multiply an arbitrary vector:

$$[x \quad y \quad z]\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = [x \quad y \quad z]\begin{bmatrix} 2x - y \\ -x + 2y - z \\ -y + 2z \end{bmatrix}$$

$$= 2x^2 - xy - xy + 2y^2 - yz - yz + 2z^2 = 2x^2 - 2xy + 2y^2 - 2yz + 2z^2$$

$$= x^2 + (x^2 - 2x + y^2) + (y^2 - 2yz + z^2) + z^2$$

$$\boxed{= x^2 + (x - y)^2 + (y - z)^2 + z^2}$$

This is never negative, so the matrix is positive definite

Most matrices $\boldsymbol{A}$ can be written as the product of two matrices, $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$, where $\boldsymbol{L}$ is lower-triangular and $\boldsymbol{U}$ is upper triangular

- This is called the *LU decomposition* of A
- For example,

$$A = \begin{bmatrix} 4 & 2 & 1 \\ 5 & 1 & 3 \\ 2 & 2 & 1 \end{bmatrix}, \quad A = LU = \begin{bmatrix} 1 & 0 & 0 \\ 0.4 & 1 & 0 \\ 0.8 & 0.75 & 1 \end{bmatrix} \begin{bmatrix} 5 & 1 & 3 \\ 0 & 1.6 & -0.2 \\ 0 & 0 & -1.25 \end{bmatrix}$$

$\boldsymbol{LU}$ can be used to solve the matrix equation $\boldsymbol{A}x = b$

- Consider $\boldsymbol{A}x = b$; find the $\boldsymbol{LU}$ decomposition of $\boldsymbol{A}$

- $\boldsymbol{A}x = b \;\Rightarrow \boldsymbol{LU}x = b \Rightarrow \boldsymbol{L}(\boldsymbol{U}x) = b$ let $\boldsymbol{U}x = y$
- Solve $\boldsymbol{L}y = b$ for y; then $\boldsymbol{U}x = y$ can be solved for x
- This is advantageous since L and U are triangular and their solutions are simple

$$Ly = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$y_1 = \frac{b_1}{l_{11}} \qquad y_2 = \frac{b_2 - l_{21} y_1}{l_{22}} \qquad y_3 = \frac{b_3 - l_{32} y_2 - l_{31} y_1}{l_{33}}$$

Example of solving a set of linear equations using LU decomposition

- $p + 2q + 3r = 5$
- $2p - 4q + 6r = 18$
- $3p - 9q - 3r = 6$

$$Ax = b: \begin{bmatrix} 1 & 2 & 3 \\ 2 & -4 & 6 \\ 3 & -9 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 18 \\ 6 \end{bmatrix}$$

find the $\boldsymbol{LU}$ decomposition of $\boldsymbol{A}$

- $\boldsymbol{A} = \boldsymbol{LU}$ $\quad \begin{bmatrix} 1 & 2 & 3 \\ 2 & -4 & 6 \\ 3 & -9 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & -8 & 0 \\ 3 & -15 & -12 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
- $\boldsymbol{L}y = b = \begin{bmatrix} 1 & 0 & 0 \\ 2 & -8 & 0 \\ 3 & -15 & -12 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 18 \\ 6 \end{bmatrix};$ math $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$
- Now, $\boldsymbol{U}x = y = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix};$ math $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$

There are a few important facts about LU decomposition

- If <u>any</u> element on the diagonal is zero, then the matrix either has no LU decomposition, or (if the matrix is also singular), there are infinitely many possible LU decompositions

- LU decomposition is <u>not</u> unique, in general – there can be more than one valid decomposition for a matrix A

VIRGINIA TECH.
ece
BRADLEY DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

Solve this set of linear equations using LU decomposition

- $2x + 4y - z = 2$
- $3x + 2y + 2z = 1$
- $4x + 2y + z = 0$

$$Ax = b:\quad \begin{bmatrix} 2 & 4 & -1 \\ 3 & 2 & 2 \\ 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

find the $\boldsymbol{LU}$ decomposition of $\boldsymbol{A}$

- $\boldsymbol{A} = \boldsymbol{LU}$ $\quad \begin{bmatrix} 2 & 4 & -1 \\ 3 & 2 & 2 \\ 4 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 0.75 & 0.167 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 & 1 \\ 0 & 3 & -1.5 \\ 0 & 0 & 1.5 \end{bmatrix}$
- $\boldsymbol{L}y = b = \begin{bmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 0.75 & 0.167 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix};$ math $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1.5 \end{bmatrix}$
- Now, $\boldsymbol{U}x = y = \begin{bmatrix} 4 & 2 & 1 \\ 0 & 3 & -1.5 \\ 0 & 0 & 1.5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1.5 \end{bmatrix};$ math $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -0.556 \\ 0.889 \\ 0.444 \end{bmatrix}$

There are a number of other forms of decomposition for matrices; some exist even when an LU decomposition does not

- $\boldsymbol{LUP}$ factorization: $\boldsymbol{A} = \boldsymbol{LUP}$, where $\boldsymbol{L}$ is lower-diagonal, $\boldsymbol{U}$ is upper diagonal and $\boldsymbol{P}$ is a permutation matrix
  - For example, $P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

- $\boldsymbol{LDU}$ factorization: $\boldsymbol{A} = \boldsymbol{LDU}$, where $\boldsymbol{L}$ is lower-diagonal, $\boldsymbol{U}$ is upper diagonal and $\boldsymbol{D}$ is a diagonal matrix

- Eigendecomposition – based on eigenvectors
  - Also called spectral decomposition
  - Very important; see next lecture

Remember that $n \times n$ matrices define a transformation that will scale and offset a vector

- Consider the following matrix transformation: $\boldsymbol{A} = \begin{bmatrix} -1 & 2 & 3 \\ 2 & 3 & 1 \\ 2 & 2 & 2 \end{bmatrix}$
- What if we apply this to the vector $v = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$?
- $\begin{bmatrix} -1 & 2 & 3 \\ 2 & 3 & 1 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} -1/3 - 4/3 + 2 \\ 2/3 - 2 + 2/3 \\ 2/3 - 4/3 + 4/3 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$
- Multiplication by the matrix $\boldsymbol{A}$ leaves this vector v unchanged

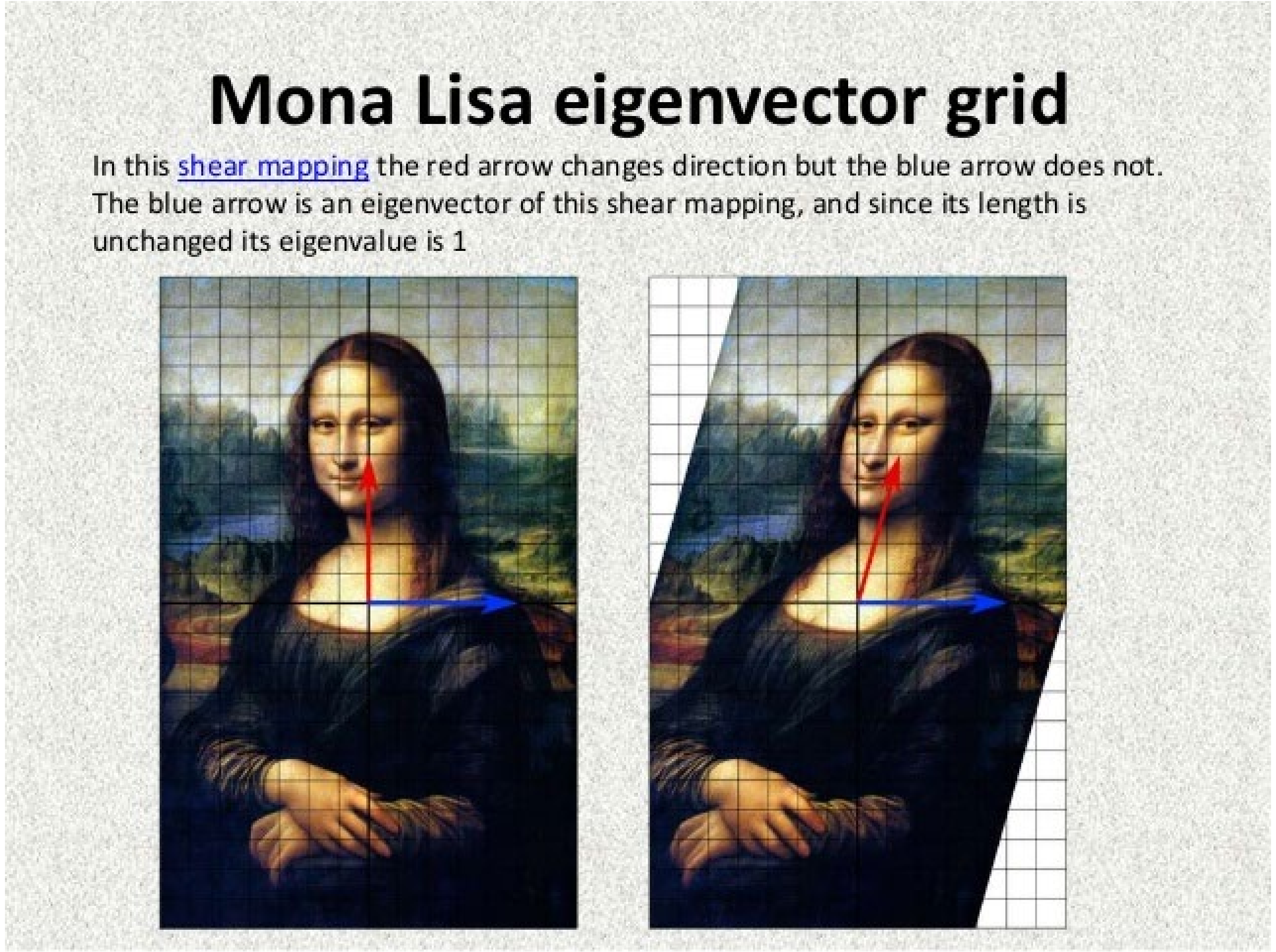
For a given matrix $\boldsymbol{A}$, a vector v that is unchanged in direction by premultiplication by $\boldsymbol{A}$ is an *eigenvector* of $\boldsymbol{A}$

VIRGINIA TECH.

ece BRADLEY DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

- In this example $\begin{bmatrix} -1 & 2 & 3 \\ 2 & 3 & 1 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$
- In general there may be a scaling constant
  - The vector points in the same direction but may have different magnitude
- In other words, $\boldsymbol{A}v = \lambda v$
- Generally, an $n \times n$ matrix will have n eigenvectors with n corresponding constants λ – called eigenvalues

Transformation by the matrix $\boldsymbol{A}$ will not rotate its eigenvectors…



The characteristic equation is often written as
$(\boldsymbol{A} - \lambda \boldsymbol{I})v = 0$

- We find the eigenvalues first

- We then have a set of matrix equations that can each be solved for one of the eigenvectors

- If $(\boldsymbol{A} - \lambda \boldsymbol{I})v = 0$, then $det(\boldsymbol{A} - \lambda \boldsymbol{I}) = \mathbf{0}$
- This yields the characteristic equation for λ

$$A = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}; \; det(\boldsymbol{A} - \lambda \boldsymbol{I}) = det\left(\begin{bmatrix} 3-\lambda & 1 \\ 2 & 4-\lambda \end{bmatrix}\right) = \mathbf{0}$$

$$det(\boldsymbol{A} - \lambda \boldsymbol{I}) = (\mathbf{3} - \boldsymbol{\lambda})(\mathbf{4} - \boldsymbol{\lambda}) - \mathbf{2} = \boldsymbol{\lambda}^{\mathbf{2}} - \mathbf{7}\boldsymbol{\lambda} + \mathbf{10}$$

$$\lambda_1 = \frac{7+\sqrt{49-40}}{2} = \frac{7+\sqrt{9}}{2} = 5 \;\; ; \; \lambda_2 = \frac{7-\sqrt{49-40}}{2} = \frac{7-\sqrt{9}}{2} = 2$$

For a given eigenvalue λ, we can solve the equation $(\boldsymbol{A} - \lambda \boldsymbol{I})v = 0$ to find the corresponding eigenvector v

$$\lambda_1 = 5 \quad ; \; \lambda_2 = 2$$

$$(\boldsymbol{A} - \lambda \boldsymbol{I})v = \begin{bmatrix} 3-\lambda & 1 \\ 2 & 4-\lambda \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$(3-\lambda)v_1 + v_2 = 0 \qquad 2v_1 + (4-\lambda)v_2 = 0$$

Note: there can be many solutions to these equations, differing in scale:

$$\lambda_1 = 2 : \quad v = \begin{bmatrix} -0.707 \\ 0.707 \end{bmatrix}$$

$$\lambda_1 = 5 : \quad v = \begin{bmatrix} -0.447 \\ -0.894 \end{bmatrix}$$

An example: $A = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 3 & 1 & -3 \end{bmatrix}$

$$\begin{vmatrix} 1-\lambda & 0 & 4 \\ 0 & 2-\lambda & 0 \\ 3 & 1 & -3-\lambda \end{vmatrix} = (1-\lambda)(\lambda^2+\lambda-7)+4(6-3\lambda)$$

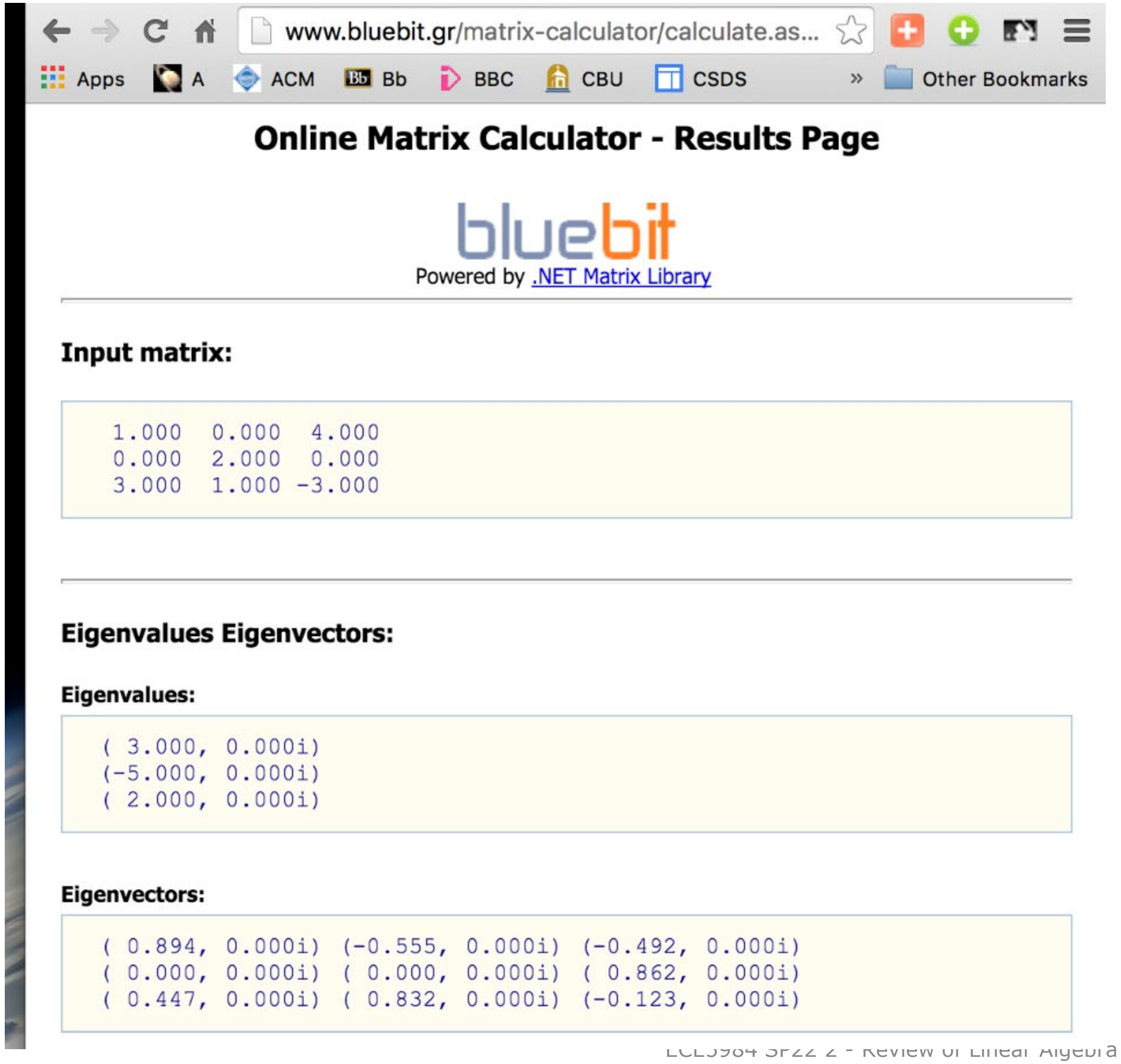
$$= -\lambda^3 + 18\lambda - 9 \qquad \lambda_1 = 2,\ \lambda_2 = 3,\ \lambda_3 = -5$$

$$\lambda_1 = 2: \begin{bmatrix} -1 & 0 & 4 \\ 0 & 0 & 0 \\ 3 & 1 & -5 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}: \qquad \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 4 \\ -7 \\ 1 \end{bmatrix}$$

$$\lambda_2 = 3: \begin{bmatrix} -2 & 0 & 4 \\ 0 & -1 & 0 \\ 3 & 1 & -6 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}: \qquad \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

$$\lambda_3 = -5: \begin{bmatrix} 6 & 0 & 4 \\ 0 & 7 & 0 \\ 3 & 1 & 4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}: \qquad \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -3 \end{bmatrix}$$

Matrix calculation gives us the same result – or does it?



Different eigenvector solution methods may find vectors that are linear factors of others

$$\begin{bmatrix} 0.894 \\ 0 \\ 0.447 \end{bmatrix} = k \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} -0.555 \\ 0 \\ 0.832 \end{bmatrix} = l \begin{bmatrix} 2 \\ 0 \\ -3 \end{bmatrix}$$

$$\begin{bmatrix} -0.492 \\ 0.862 \\ -0.123 \end{bmatrix} = m \begin{bmatrix} 4 \\ -7 \\ 1 \end{bmatrix}$$

$A = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$ Find the eigenvalues and eigenvectors

$A = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$ Find the eigenvalues and eigenvectors

$Ax = \lambda x$ or $det(A - \lambda I) = 0$

$$|A - \lambda I| = 0 = \begin{vmatrix} 1-\lambda & 2 \\ 3 & -4-\lambda \end{vmatrix} = (1-\lambda)(-4-\lambda) - 6 = 0$$

$$-4 - \lambda + 4\lambda + \lambda^2 - 6 = \lambda^2 + 3\lambda - 10 = 0 \qquad \lambda = 2, -5$$

To find eigenvalues, $Ax = \lambda x$; $\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}\begin{bmatrix} x_{n1} \\ x_{n2} \end{bmatrix} = \lambda_n \begin{bmatrix} x_{n1} \\ x_{n2} \end{bmatrix}$

For $\lambda = -5$:

$6x_{11} + 2x_{12} = 0 \quad 3x_{11} + x_{12} = 0$ choose $x_{11} = 1$: $x_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$

For $\lambda = 2$:

$-x_{21} + 2x_{22} = 0 \quad 3x_{21} - 6x_{22} = 0$ choose $x_{22} = 1$: $x_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

Real-valued matrices can have complex eigenvalues under some circumstances – and eigenvectors may not exist even when eigenvalues do (?)

Matrix Type	Eigendecomposition Properties
Square Symmetric	**Eigenvalues**: always real, nonnegative **Eigenvectors**: always orthogonal
Square Asymmetric	**Eigenvalues**: can be complex **Eigenvectors**: don't necessarily exist
Non-square	Eigendecomposition not possible

The Cayley-Hamilton theorem says that any square matrix $\boldsymbol{A}$ satisfies its own characteristic equation

$$A = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix};$$

$$det(\boldsymbol{A} - \lambda \boldsymbol{I}) = \boldsymbol{\lambda^2} - 7\boldsymbol{\lambda} + 10 = 0$$

Now, apply Cayley-Hamilton:

$$\boldsymbol{A^2} - 7\boldsymbol{A} + 10 = \boldsymbol{AA} - 7\boldsymbol{A} + 10\boldsymbol{I}$$

$$\begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix} - 7\begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix} + 10\boldsymbol{I} = \begin{bmatrix} 11 & 7 \\ 14 & 18 \end{bmatrix} - \begin{bmatrix} 21 & 7 \\ 14 & 28 \end{bmatrix} + 10\boldsymbol{I}$$

$$= \begin{bmatrix} -10 & 0 \\ 0 & -10 \end{bmatrix} + 10\boldsymbol{I} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}$$

For an $n \times n$ square matrix, there are *n* eigenvalues – and they satisfy some properties

- One or more eigenvalues may be duplicated, if the characteristic equation has multiple roots with the same value
- A real-valued matrix may have complex eigenvalues – if so, they will occur in complex conjugate pairs

- The product of the n eigenvalues equals the determinant.
- The sum of the n eigenvalues equals the trace of the matrix (the sum of the diagonal elements)

The *spectral matrix* of a square matrix $\boldsymbol{A}$ is the matrix with the eigenvalues of $\boldsymbol{A}$ on the diagonal and all other entries zero

- If $\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 3 & 1 & -3 \end{bmatrix}$ then the spectral matrix $\boldsymbol{S} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -5 \end{bmatrix}$, because the eigenvalues of $\boldsymbol{A}$ are 2, 3 and -5.

- The multiset of the eigenvalues is called the *spectrum* of the matrix $\boldsymbol{A}$.

The *modal matrix* of a square matrix $\boldsymbol{A}$ is the matrix formed by concatenating the eigenvectors

- If $\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 3 & 1 & -3 \end{bmatrix}$, eigenvectors $v_1 - v_3$ are $\begin{bmatrix} 4 \\ -7 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -3 \end{bmatrix}$.

- The modal matrix of $\boldsymbol{A}$ is then $\boldsymbol{M} = \begin{bmatrix} 4 & 2 & 2 \\ -7 & 0 & 0 \\ 1 & 1 & -3 \end{bmatrix}$.
- Pre- and post-multiplying $\boldsymbol{A}$ by $\boldsymbol{M}^{-1}$and $\boldsymbol{M}$ produces the spectral matrix $\boldsymbol{S}$ (which is diagonal)…

- $\boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M} = \begin{bmatrix} 0 & {}^{-1}/_{7} & 0 \\ {}^{3}/_{8} & {}^{1}/_{4} & {}^{1}/_{4} \\ {}^{1}/_{8} & {}^{1}/_{28} & {}^{-1}/_{4} \end{bmatrix} \begin{bmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 3 & 1 & -3 \end{bmatrix} \begin{bmatrix} 4 & 2 & 2 \\ -7 & 0 & 0 \\ 1 & 1 & -3 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -5 \end{bmatrix}$

Eigendecomposition forms the basis of *factor analysis* – a tool for understanding multivariate data

- Imagine a set of N measurements in D multiple variables
- We can form the mean vector and the covariance matrix:
- $\mu = \frac{1}{N}\begin{bmatrix} \sum_{i=1}^{N} x_{1i} \\ \sum_{i=1}^{N} x_{2i} \\ \vdots \\ \sum_{i=1}^{N} x_{Di} \end{bmatrix},$
- $\Sigma = \frac{1}{N}\begin{bmatrix} \sum_{i=1}^{N}(x_1-\mu_1)(x_1-\mu_1) & \sum_{i=1}^{N}(x_1-\mu_1)(x_2-\mu_2) & \cdots & \sum_{i=1}^{N}(x_1-\mu_1)(x_D-\mu_D) \\ \sum_{i=1}^{N}(x_2-\mu_2)(x_1-\mu_1) & \sum_{i=1}^{N}(x_2-\mu_2)(x_2-\mu_2) & \cdots & \sum_{i=1}^{N}(x_2-\mu_2)(x_D-\mu_D) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{N}(x_D-\mu_D)(x_1-\mu_1) & \sum_{i=1}^{N}(x_D-\mu_D)(x_2-\mu_2) & \cdots & \sum_{i=1}^{N}(x_D-\mu_D)(x_D-\mu_D) \end{bmatrix}$

The eigenvectors of the covariance matrix define a new coordinate system, aligned with the axis of maximum variance in the data

VIRGINIA TECH.
BRADLEY DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

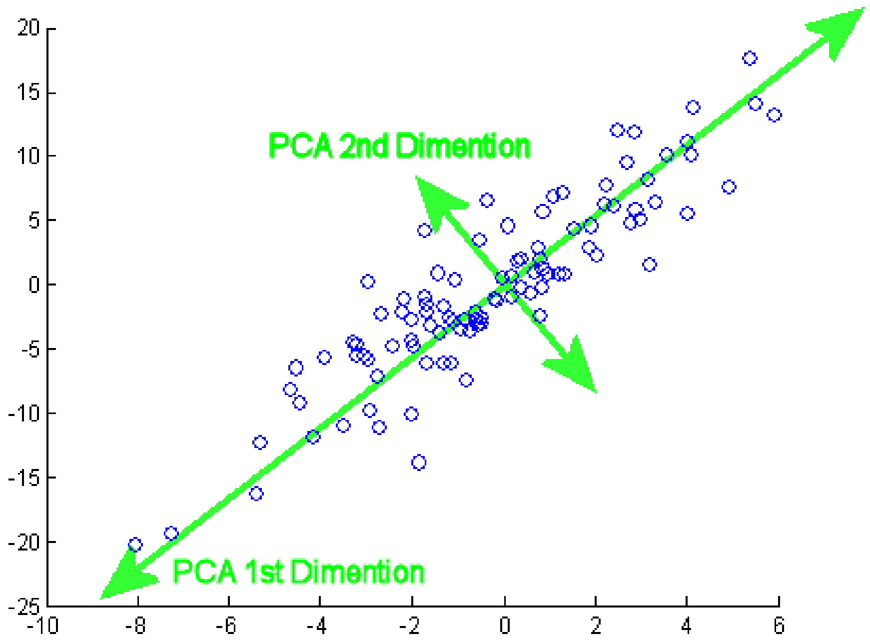
- Analysis of the covariance matrix $\boldsymbol{\Sigma}$ of the data reveals the distribution of variation in the data samples

- This is called Principal Component Analysis
  - The vector associated with the eigenvalue of highest magnitude is the axis along which the data has the highest variance
  - The portion of total data variance which occurs in that "direction" is the ratio of that eigenvector to the sum of all:

$$r = \frac{\lambda_{max}}{\lambda_1 + \lambda_2 + \cdots + \lambda_D}$$

  - The second highest eigenvector is associated with the perpendicular axis with the highest portion of the remaining variance

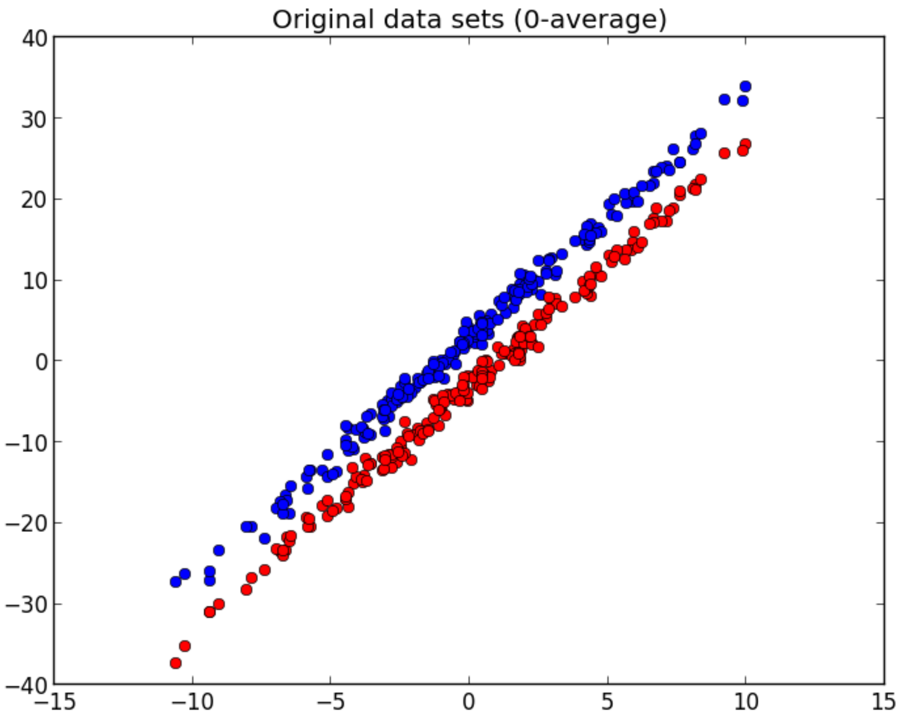
Principal Components Analysis on 2-dimensional data



- The 1st dimension clearly exhibits the most variance
- The 2nd dimension is orthogonal to it, and in higher dimensional cases, would be the next-highest variance dimension

PCA for Dimensionality reduction – don't keep all of the new axes

- By ranking the eigenvectors of $\boldsymbol{\Sigma}$ by decreasing magnitude of λ , we can keep only the most significant
- Using the new axes defined by the key eigenvectors, we can represent the data in a lower dimensional space
- The degree of "approximation" experienced can be estimated by analysis of the sequence of eigenvalues



Today’s Objectives

- Vectors
  - Vector products
  - Vector differentiation
  - Vector norms
- Matrices
  - Matrix products
  - Inversion and singularity
  - Special matrices
  - Solving systems of equations
  - Matrix decomposition
  - Eigendecomposition