Linear Algebra and Optimization

Ankur Moitra

Developed and co-taught with Pablo Parrilo

Online Seminar on Undergraduate Math Education

OUTLINE

(1) **What is it?**

Goals, Structure and Audience

(2) **How is it different?**

Pedagogy, Computation

(3) **Where did it come from?**

Math Foundations for Data Science

OVERVIEW

Goals: (1) Introduction to linear algebra and optimization, with a view towards **modeling, computation and applications**.

OVERVIEW

Goals: (1) Introduction to linear algebra and optimization, with a view towards **modeling, computation and applications**.

(2) Solid, broad foundation for engineering and science, across the Institute.

OVERVIEW

Goals: (1) Introduction to linear algebra and optimization, with a view towards **modeling, computation and applications**.

(2) Solid, broad foundation for engineering and science, across the Institute.

Collaboration between Course 6 (EECS) and Course 18 (Math)

OVERVIEW

Goals: (1) Introduction to linear algebra and optimization, with a view towards **modeling, computation and applications**.

(2) Solid, broad foundation for engineering and science, across the Institute.

Collaboration between Course 6 (EECS) and Course 18 (Math)

Audience: Sophomore level

OVERVIEW

Goals: (1) Introduction to linear algebra and optimization, with a view towards **modeling, computation and applications**.

(2) Solid, broad foundation for engineering and science, across the Institute.

Collaboration between Course 6 (EECS) and Course 18 (Math)

Audience: Sophomore level

Piloted in Fall 2020, taught every Fall since then

STRUCTURE

Weekly psets

Online checkups

Recitations

Miniprojects (in Julia)

Three tests

Part I: Working with Vectors and Matrices
- Lecture 1: A Panoramic View of Linear Algebra
- Lecture 2: The Geometry of Linear Equations
- Lecture 3: Gaussian Elimination and Applications to Circuit Analysis
- Lecture 4: Multiplying Matrices and Applications to Counting Walks
- Lecture 5: Visualizations: Projections, Reflections, Rotations and Permutations
- Lecture 6: Vector Spaces, Linear Combinations and Column/Null Spaces

Part II: Geometric Foundations
- Lecture 7: The Rank, and its Equivalent Formulations
- Lecture 8: Linear Independence, Dimension and Bases
- Lecture 9: Orthogonality and Gram-Schmidt
- Lecture 10: The Determinant and its Properties
- Lecture 11: The Matrix Inverse, Existence and Projections
- Lecture 12: Least Squares and Regularization

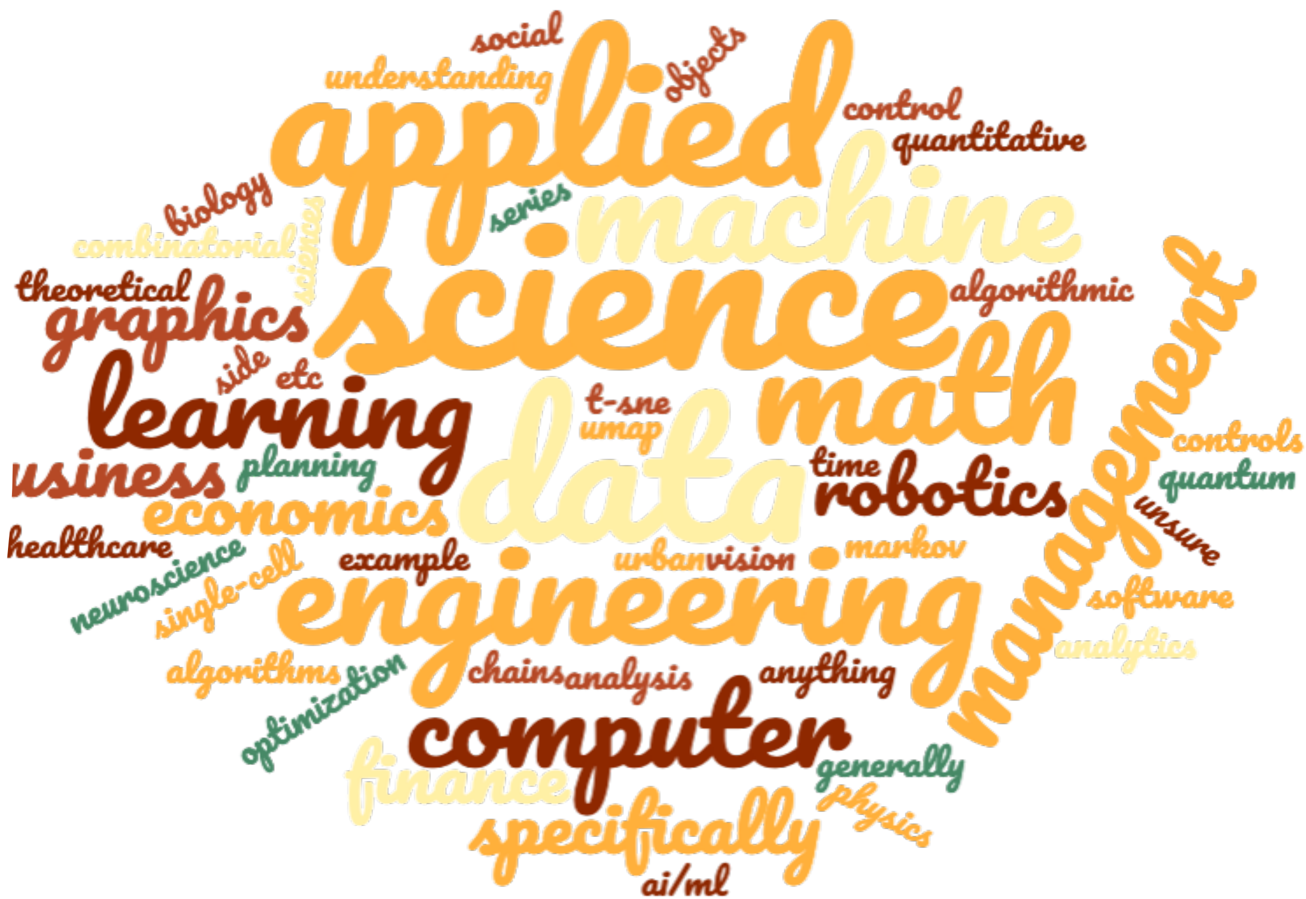
Part III: The Singular Value Decomposition and Applications
- Lecture 13: The Singular Value Decomposition
- Lecture 14: The Condition Number and Stability
- Lecture 15: Principal Component Analysis and Applications to Genetics
- Lecture 16: Word Embeddings and Exploring Biases in Data
- Lecture 17: Eigenvalues and Eigenvectors
- Lecture 18: The Eigendecomposition and Algebraic vs. Geometric Multiplicity
- Lecture 20: Markov Matrices and Applications to PageRank

Part IV: Quadratic Programming and Applications
- Lecture 21: Linear and Quadratic Programming
- Lecture 22: Support Vector Machines and the Kernel Trick
- Lecture 23: The Perceptron Algorithm

Part V: Convex Optimization and Gradient Descent

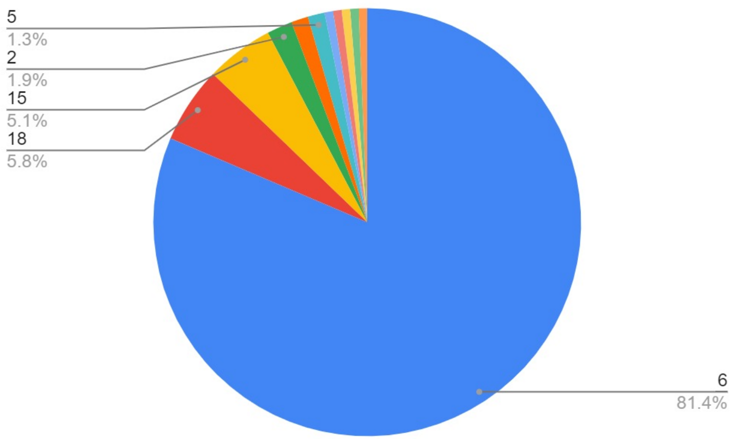
AUDIENCE

Draw on a variety of applications, already seems to appeal to a broader audience, self-reported interests:



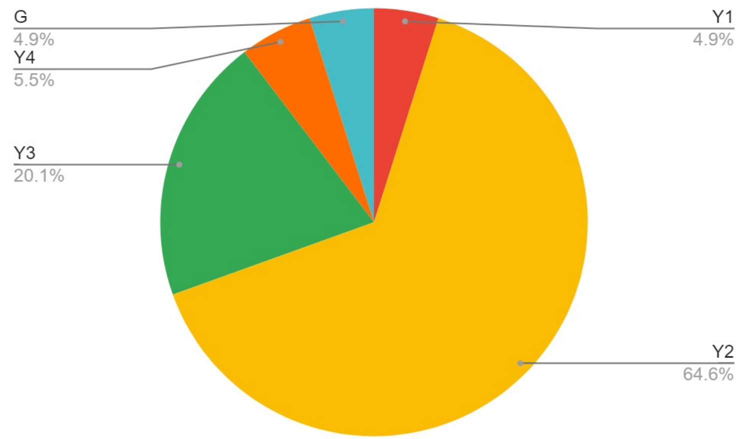
ENROLLMENT

Home Departments



	dept	registrations		dept	registrations	
EECS	6	127		14	1	Economics
Math	18	9		16	1	AeroAstro
Management	15	8		24	1	Linguistics
MechE	2	3		8	1	Physics
Urban Studies	11	2		9	1	Cog. Sci.
Chemistry	5	2				

Student Years



year	registrations
Y1	8
Y2	106
Y3	33
Y4	9
G	8

WHAT'S NEW?

Many introductory linear algebra classes are recipe-driven

WHAT'S NEW?

Many introductory linear algebra classes are recipe-driven

Can you solve this linear system?

WHAT'S NEW?

Many introductory linear algebra classes are recipe-driven

Can you solve this linear system?

Can you put this matrix in row-echelon form?

WHAT'S NEW?

Many introductory linear algebra classes are recipe-driven

Can you solve this linear system?

Can you put this matrix in row-echelon form?

Can you compute a basis for this vector space?

WHAT'S NEW?

Many introductory linear algebra classes are recipe-driven

Can you solve this linear system?

Can you put this matrix in row-echelon form?

Can you compute a basis for this vector space?

Can you compute the eigenvalues?

WHAT'S NEW?

Solving things by hand is a first step, but not the end goal

WHAT'S NEW?

Solving things by hand is a first step, but not the end goal

Can you recognize when the problem you want to solve is actually just linear algebra in disguise?

WHAT'S NEW?

Solving things by hand is a first step, but not the end goal

Can you recognize when the problem you want to solve is actually just linear algebra in disguise?

Drawing on applications spanning science and engineering:

(1) Practice applying linear algebra concepts in the wild

WHAT'S NEW?

Solving things by hand is a first step, but not the end goal

Can you recognize when the problem you want to solve is actually just linear algebra in disguise?

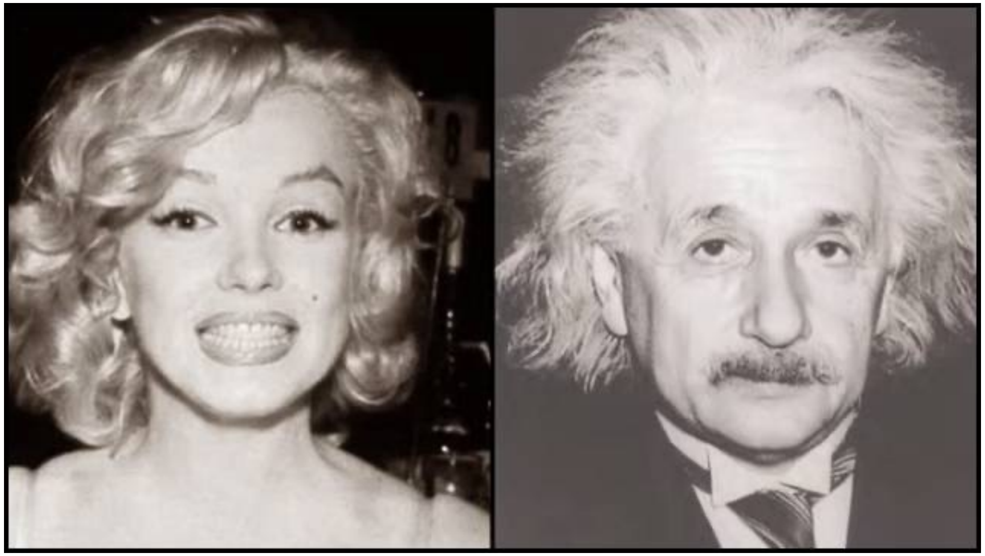
Drawing on applications spanning science and engineering:

(1) Practice applying linear algebra concepts in the wild

(2) Build familiarity with computational tools (i.e. Julia)

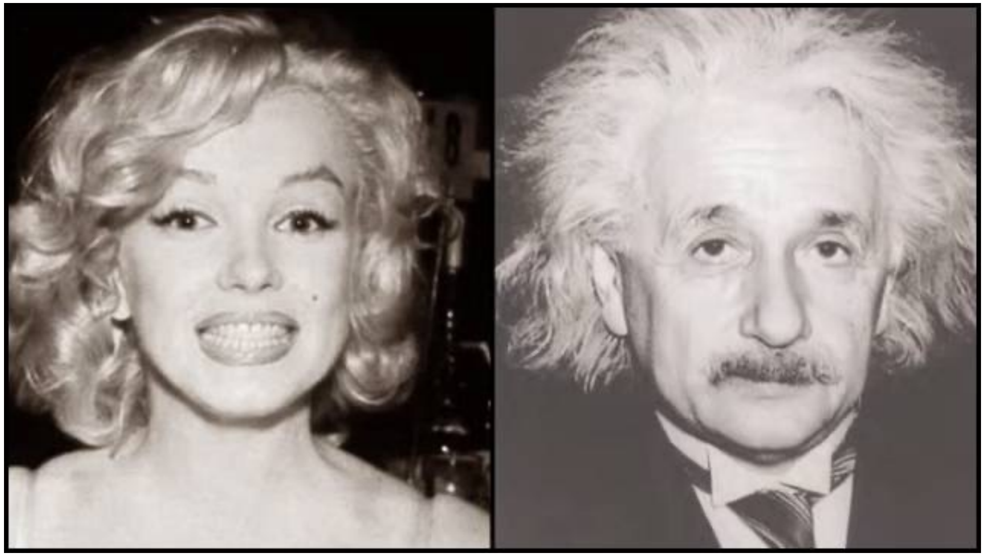
WHAT IS A VECTOR AND WHY DO WE CARE?

Powerful way to represent and manipulate data, e.g.



WHAT IS A VECTOR AND WHY DO WE CARE?

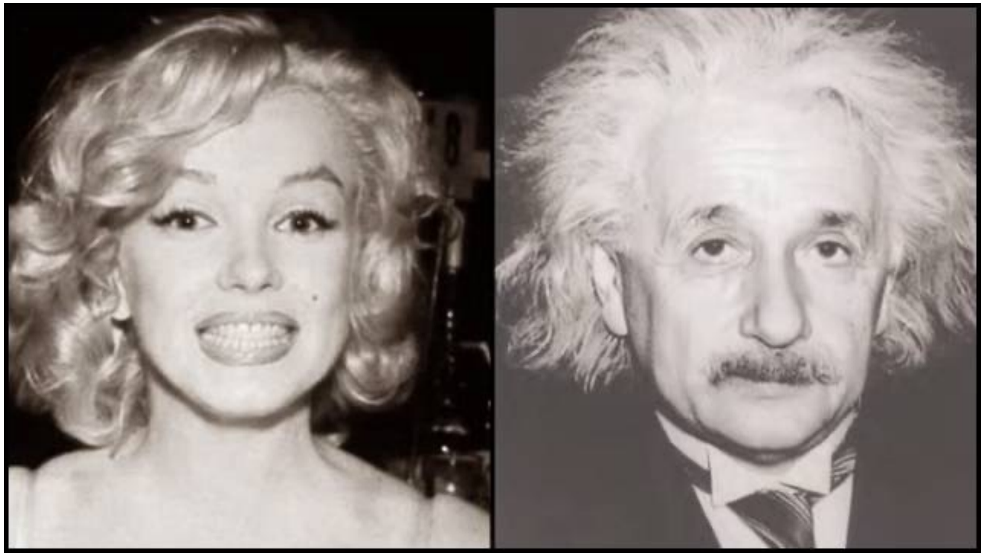
Powerful way to represent and manipulate data, e.g.



Each image can be thought of as a 115200 dimensional vector

WHAT IS A VECTOR AND WHY DO WE CARE?

Powerful way to represent and manipulate data, e.g.



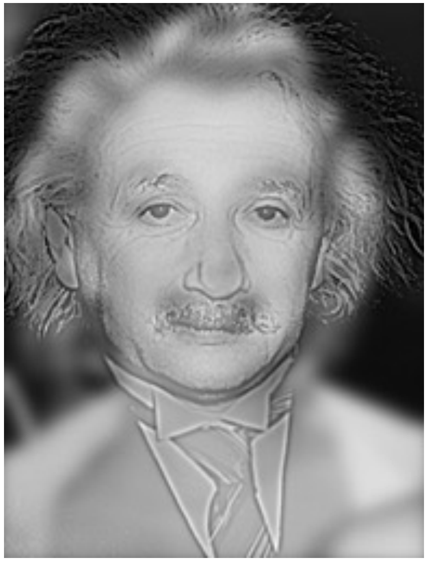
Each image can be thought of as a 115200 dimensional vector

How can we manipulate them in interesting ways?

WHAT IS A VECTOR AND WHY DO WE CARE?

Powerful way to represent and manipulate data, e.g.

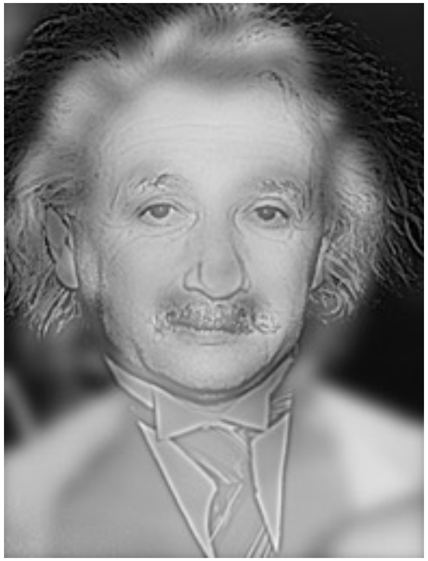
Composite image,
[Oliva, Torralba]



WHAT IS A VECTOR AND WHY DO WE CARE?

Powerful way to represent and manipulate data, e.g.

Composite image,
[Oliva, Torralba]



How do we represent these operations (blur, find edges)?

VISUAL EXAMPLES

Why is the box rotating?

VISUAL EXAMPLES

Why is the box rotating?

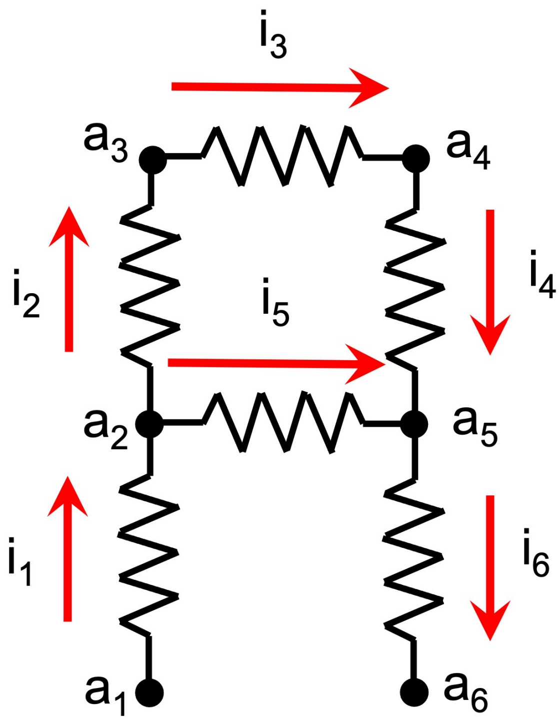
Composing two reflections gives a rotation. What is the angle?

LINEAR ALGEBRA IN DISGUISE

Nature solves linear algebra problems all the time

LINEAR ALGEBRA IN DISGUISE

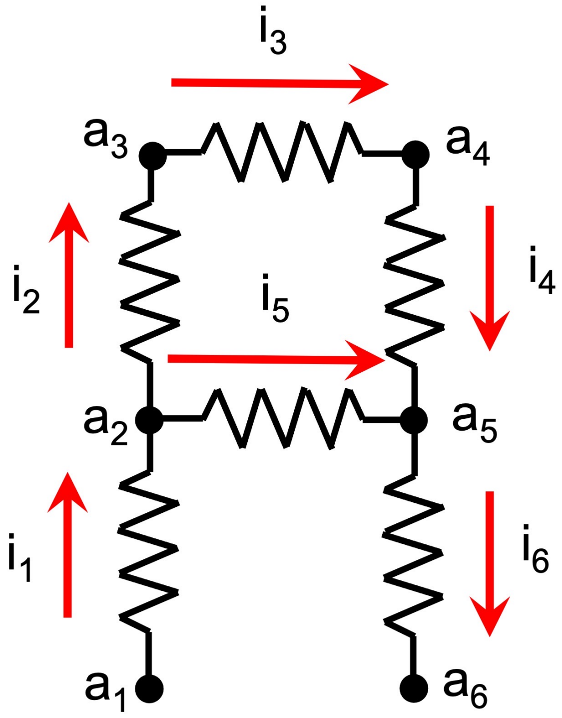
Nature solves linear algebra problems all the time



If we send one unit of current in through a_1 and out of a_6, what are the voltages at all the junctions?

LINEAR ALGEBRA IN DISGUISE

Nature solves linear algebra problems all the time



If we send one unit of current in through a_1 and out of a_6, what are the voltages at all the junctions?

Can we use Kirchoff's laws to translate this into a linear system?

PUTTING THE SVD FRONT AND CENTER

Why don't all introductory courses teach the SVD?

PUTTING THE SVD FRONT AND CENTER

Why don't all introductory courses teach the SVD?

- **More abstract, e.g. can't compute anything by hand**

PUTTING THE SVD FRONT AND CENTER

Why don't all introductory courses teach the SVD?

— **More abstract, e.g. can't compute anything by hand**

\+ **It's the Swiss army knife of linear algebra**

PUTTING THE SVD FRONT AND CENTER

Why don't all introductory courses teach the SVD?

— **More abstract, e.g. can't compute anything by hand**

\+ **It's the Swiss army knife of linear algebra**

\+ **Leads to more robust notions, e.g. approximate rank**

CONNECTIONS TO DATA SCIENCE

It's also the first thing you try in applications

CONNECTIONS TO DATA SCIENCE

It's also the first thing you try in applications

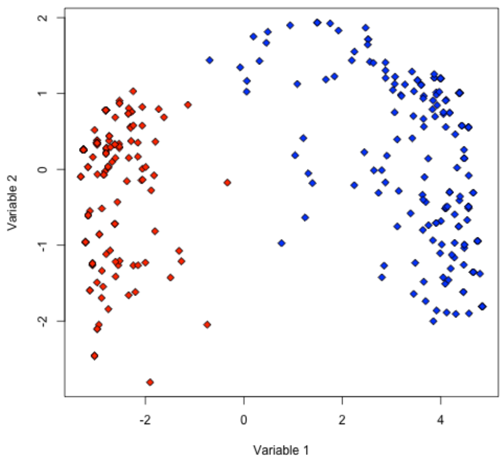
e.g. finding the largest directions of variance, to visualize high-dimensional data

CONNECTIONS TO DATA SCIENCE

It's also the first thing you try in applications

e.g. finding the largest directions of variance, to visualize high-dimensional data

US Senator Voting Records

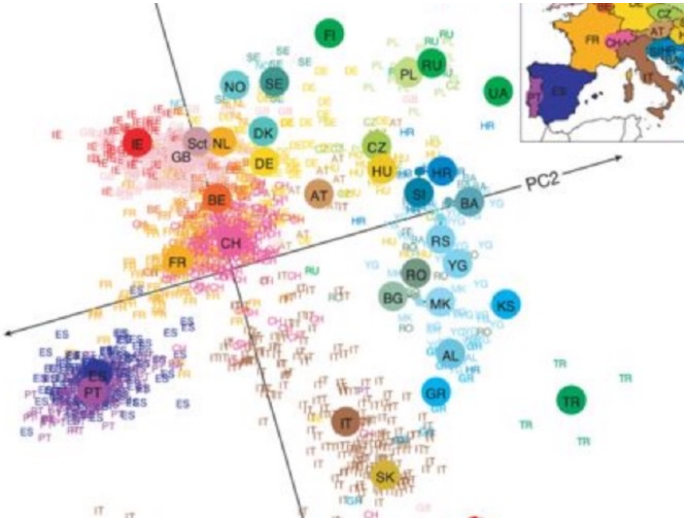


CONNECTIONS TO DATA SCIENCE

It’s also the first thing you try in applications

e.g. finding the largest directions of variance, to visualize high-dimensional data

Genes Mirror Geography



CONNECTIONS TO DATA SCIENCE

Can even understand biases in data, e.g. in word embeddings

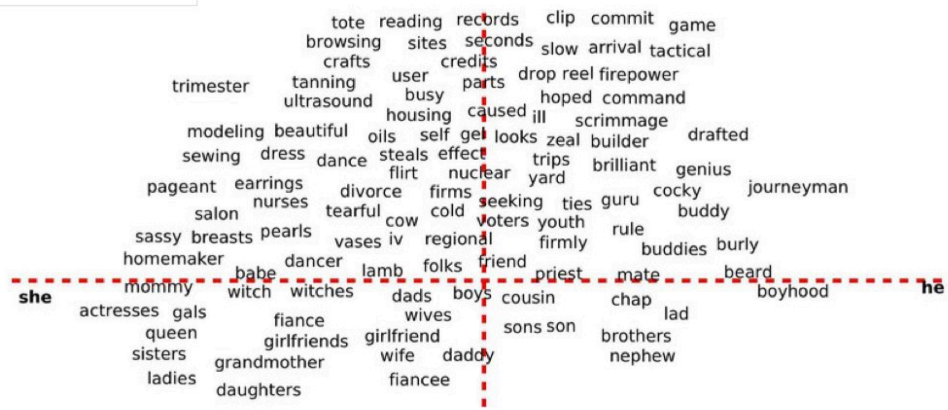
npr

INNOVATION

He's Brilliant, She's Lovely: Teaching Computers To Be Less Sexist

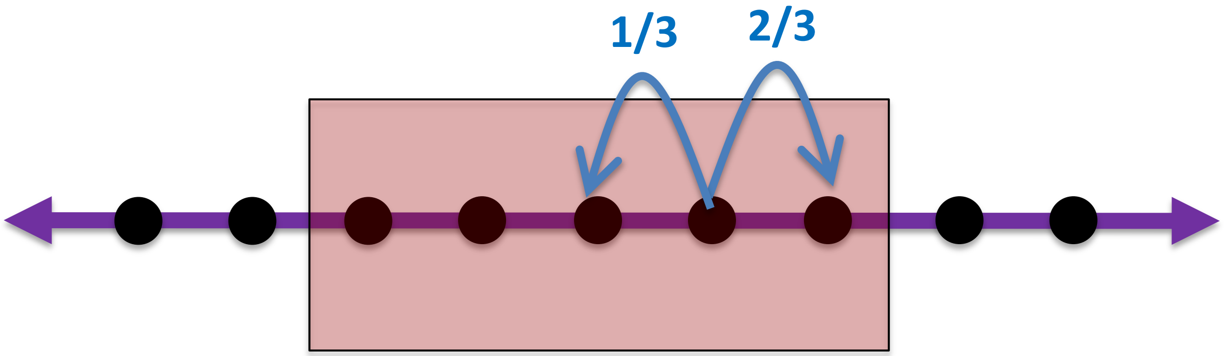
August 12, 2016 · 8:01 AM ET

BYRD PINKERTON



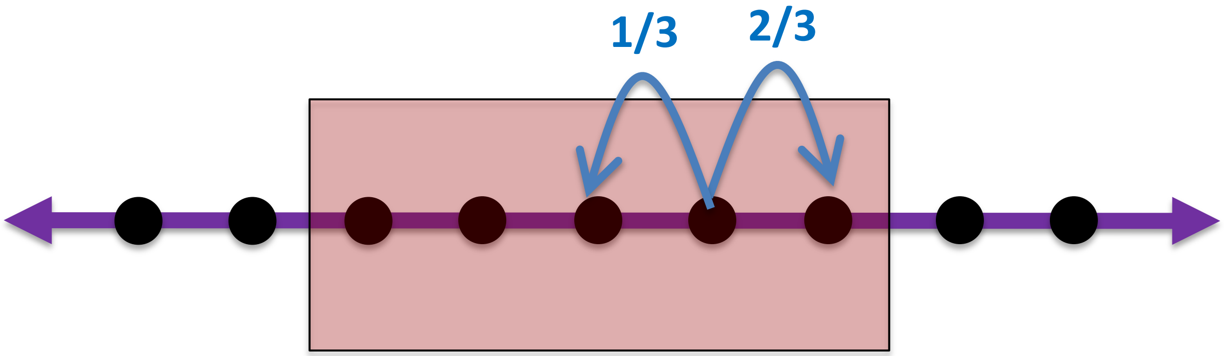
HOPPING BUNNY PROJECT

In the barrier, bunny more likely to hop outwards



HOPPING BUNNY PROJECT

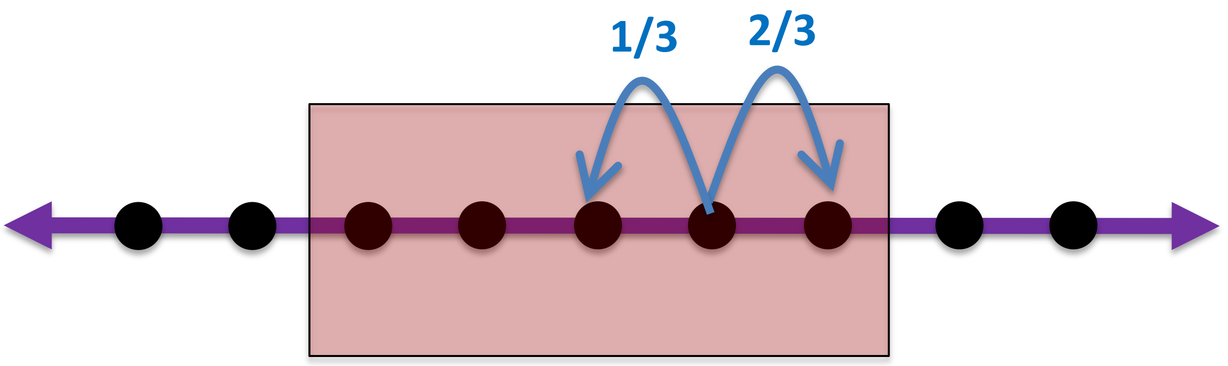
In the barrier, bunny more likely to hop outwards



How does the eigengap depend on the length of the barrier?

HOPPING BUNNY PROJECT

In the barrier, bunny more likely to hop outwards



How does the eigengap depend on the length of the barrier?

Students compute **how the coefficients in the eigenbasis change**, better behaved progress measure than distance to steady state

MENS ET MANUS

We want students to think of computational tools as a resource, e.g. for trying things out and reinforcing what they've learned

MENS ET MANUS

We want students to think of computational tools as a resource, e.g. for trying things out and reinforcing what they've learned

Lesson: Need programming to come in early, otherwise students avoid it

MORE LESSONS

Difficult to do proofs without really doing proofs, e.g.

P7. [5+5 pts] Suppose that there are square matrices A, B and T that satisfy

$$AT + B = 0.$$

(a) Show that $C(A) \subseteq C(B)$ or find a counter example

(b) Show that $C(B) \subseteq C(A)$ or find a counter example

MORE LESSONS

Difficult to do proofs without really doing proofs, e.g.

P7. [5+5 pts] Suppose that there are square matrices A, B and T that satisfy

$$AT + B = 0.$$

(a) Show that $C(A) \subseteq C(B)$ or find a counter example

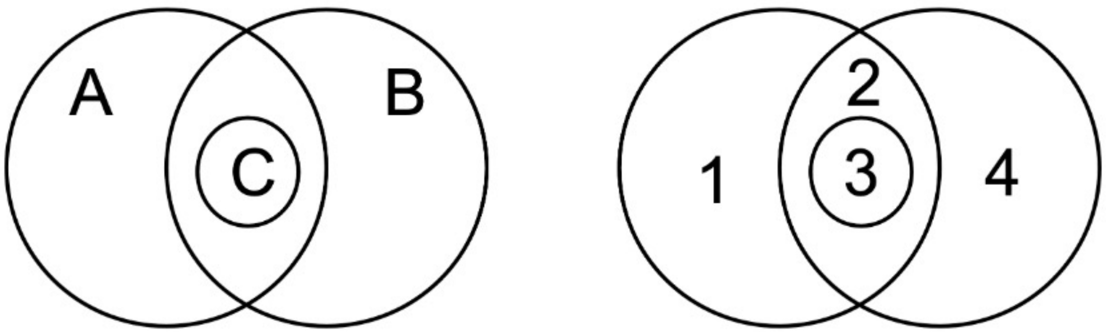
(b) Show that $C(B) \subseteq C(A)$ or find a counter example

Students find it challenging knowing how to apply a definition

MORE LESSONS

Tough to find new, creative examples of applying linear algebra year after year

P12. [5+4+3 pts] Consider the following Venn diagram:



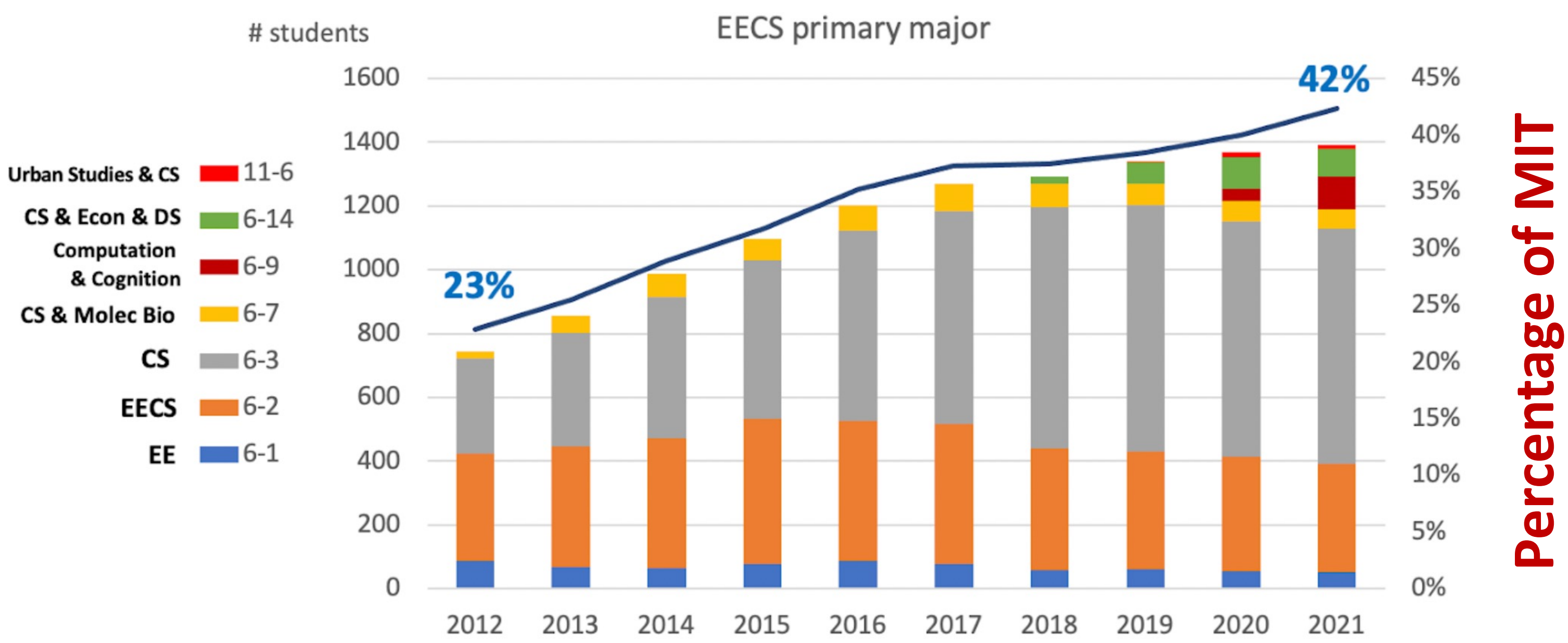
We have three sets A, B and C and there are four regions. E.g. region 2 is the set of all elements in A and B but not in C.

The notation $|A|$ means the number of elements in A. Furthermore $A \cap B$ denotes the set of elements that are in both A and B. Now suppose we know $|A| = v_A$, $|B| = v_B$, $|C| = v_C$ and $|A \cap B| = v_{AB}$.

(a) Write down a linear system to solve for the number of elements in each of the four regions, x_1, x_2, x_3 and x_4 respectively, in terms of v_A, v_B, v_C and v_{AB}.

INSTITUTIONAL CONTEXT

Massive demand for EECS



INSTITUTIONAL CONTEXT

In Spring 2020, Leslie Kaelbling chaired a committee to develop a new artificial intelligence major

INSTITUTIONAL CONTEXT

In Spring 2020, Leslie Kaelbling chaired a committee to develop a new artificial intelligence major

What all math should students know?

INSTITUTIONAL CONTEXT

In Spring 2020, Leslie Kaelbling chaired a committee to develop a new artificial intelligence major

What all math should students know?

Single variable calculus, multivariable calculus, discrete math, probability, statistics, linear algebra, optimization, differential equations

INSTITUTIONAL CONTEXT

In Spring 2020, Leslie Kaelbling chaired a committee to develop a new artificial intelligence major

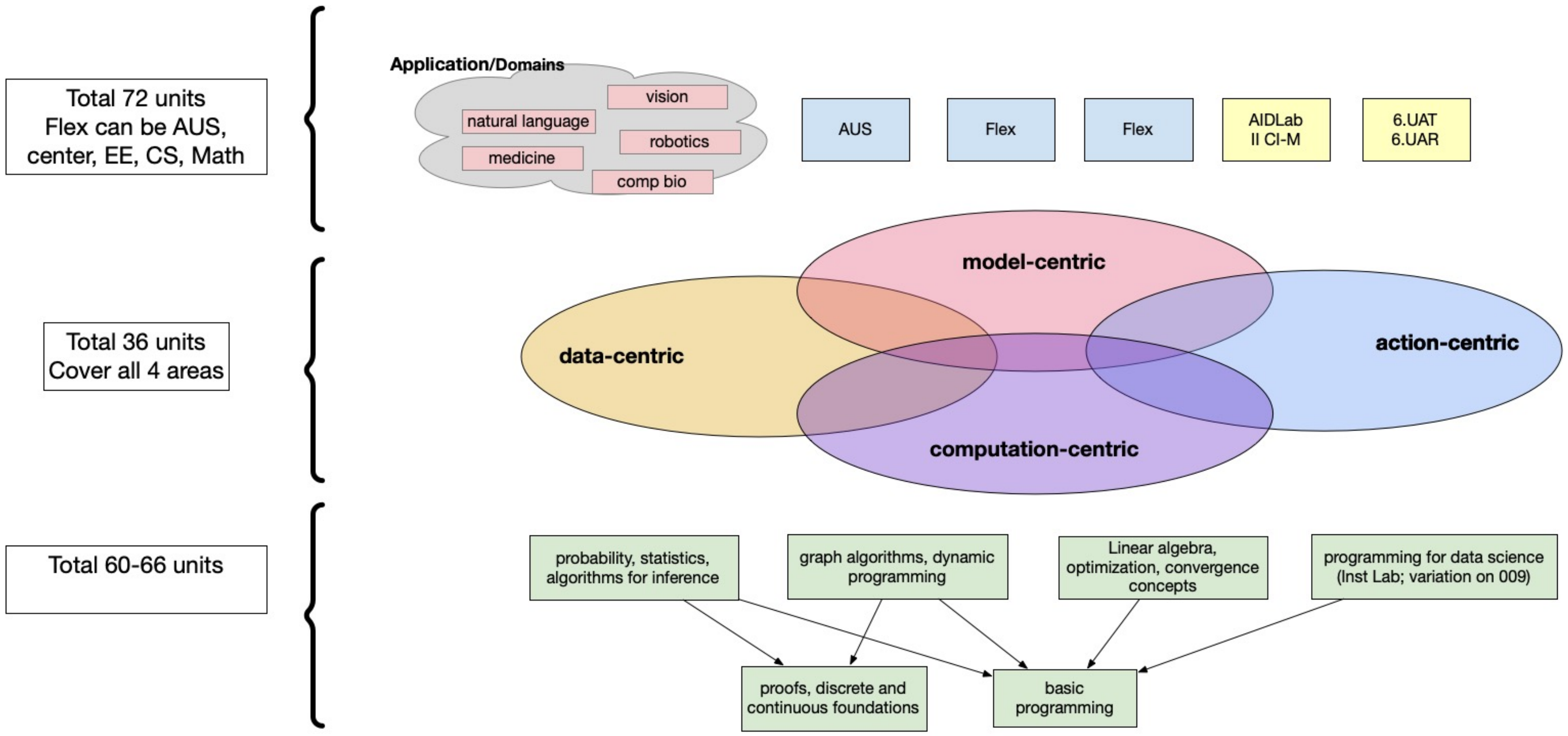
What all math should students know?

Single variable calculus, multivariable calculus, discrete math, probability, statistics, linear algebra, optimization, differential equations

Unfortunately not feasible, given other necessities, e.g. programming, machine learning, algorithms, project-based classes

INSTITUTIONAL CONTEXT

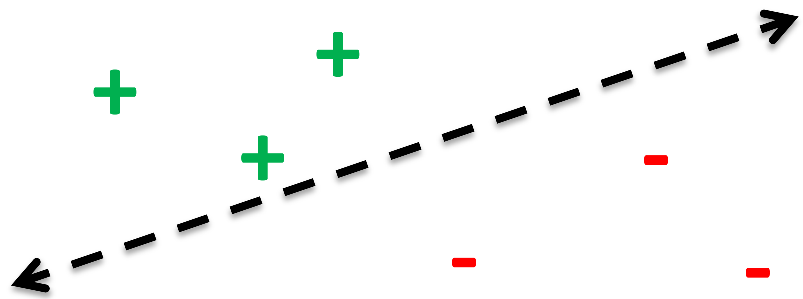
Came up with the following structure (14.5 subjects)



ANOTHER CALL TO ARMS

From interviewing faculty:

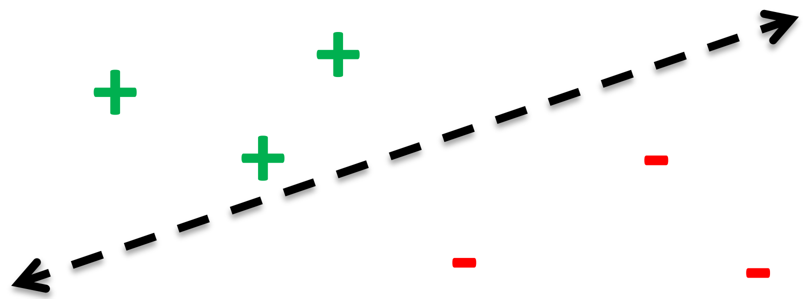
> **“Students have taken linear algebra, and yet they can’t visualize what’s happening in the perceptron algorithm”**



ANOTHER CALL TO ARMS

From interviewing faculty:

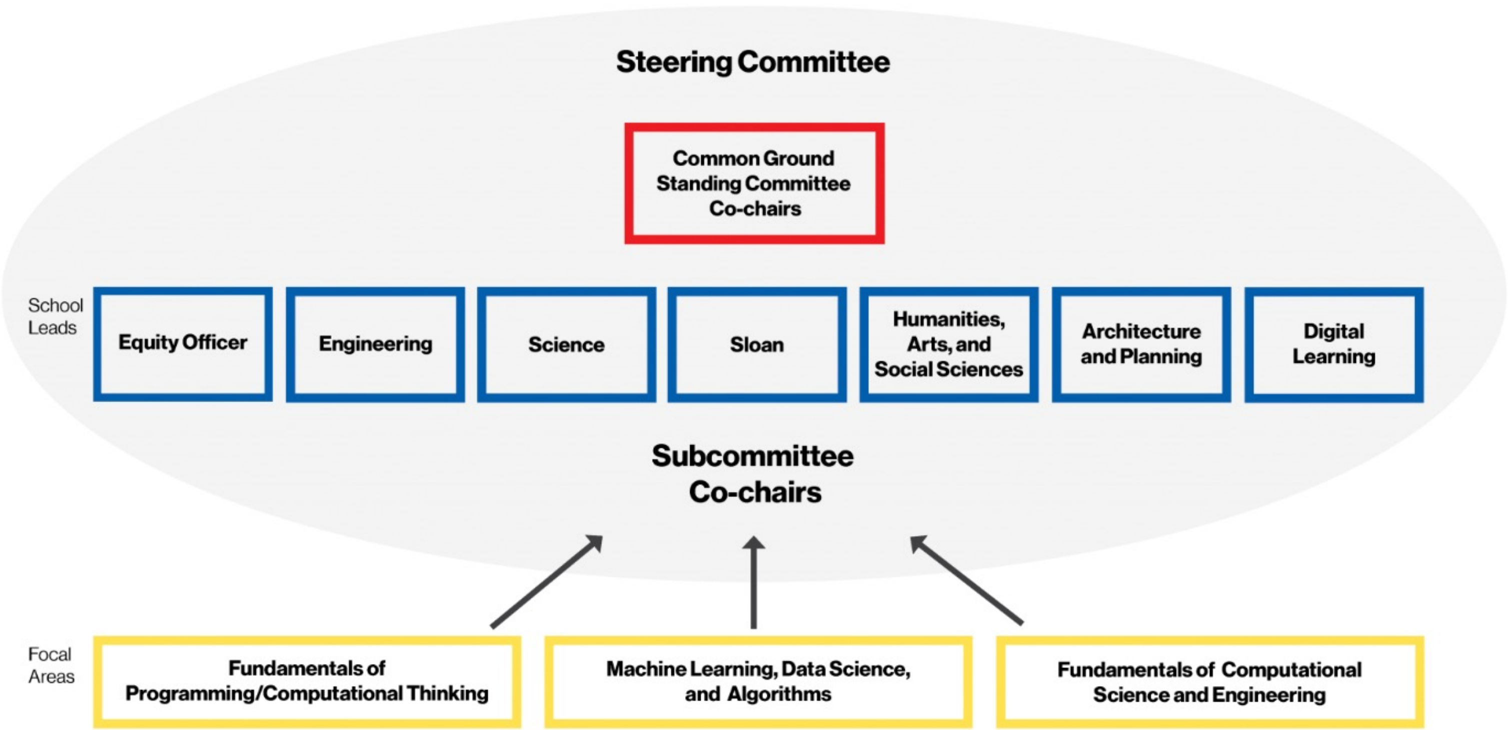
> **"Students have taken linear algebra, and yet they can't visualize what's happening in the perceptron algorithm"**



Why is that?

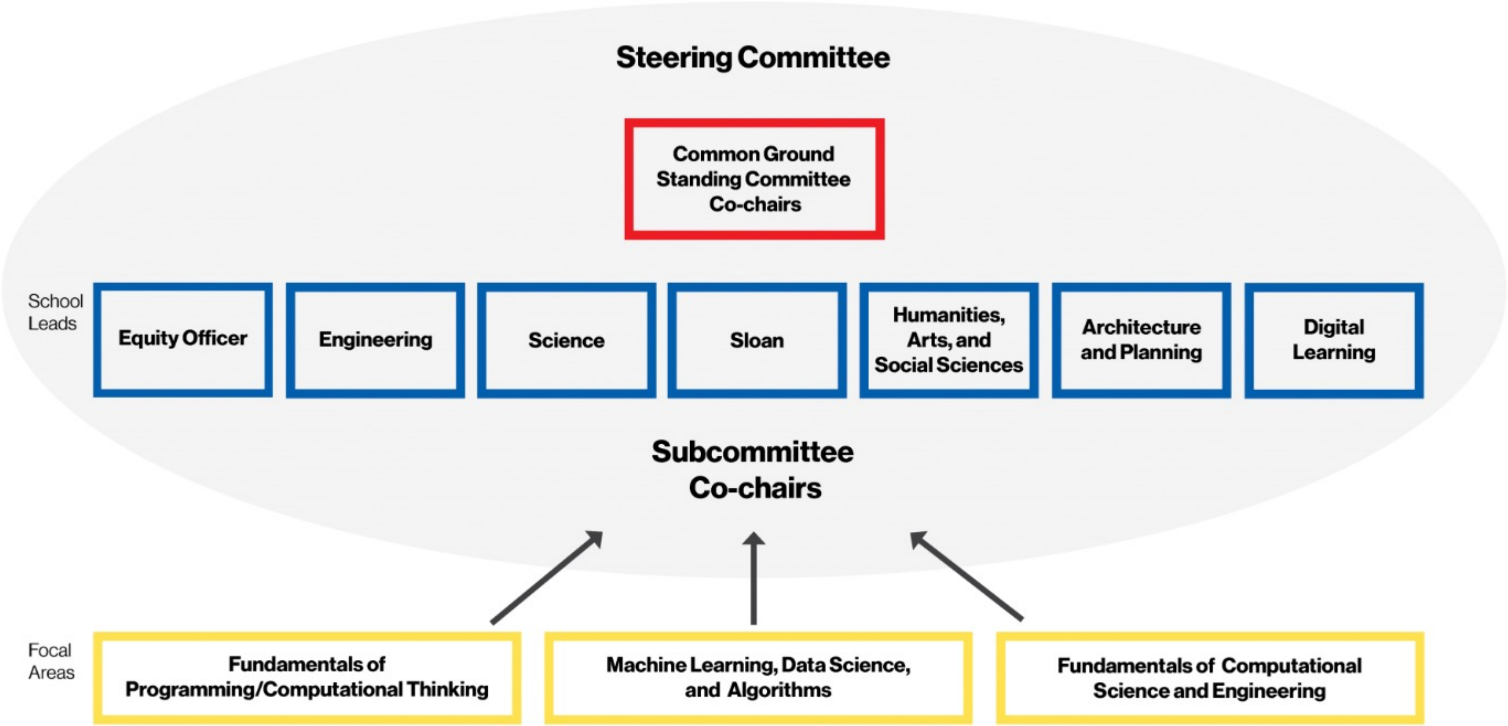
COMMON GROUND

Chaired by Asu Ozdaglar and Jeffrey Grossman



COMMON GROUND

Chaired by Asu Ozdaglar and Jeffrey Grossman



…to infuse computing education across MIT, and coordinate among departments

LOOKING FORWARD

Main takeaway:

Deliberately teach students how to make connections?

Problem → Model → Algorithm → Code → Results → Evaluation/ Interpretation

Thanks! Any Questions?