Lecture 2 18.06 (2020) Syllabus link

Is this a computer scientist's linear algebra? (of course it isn't only)

Answer: Mathematics is for everyone. Linear Algebra is for everyone.

This class will be as useful for the biologists, the economists, the physicists, the math majors, as well as the computer scientists and electrical engineers.

And maybe even a janitor or two. (for those who saw the movie "Good Will Hunting")

Summary

Today we will look at various mathematical structures that allow for linear combinations with real numbers. Informally, we will say that any such structure, discrete or continuous, will be called a vector space.

We will see that traditional 1) column vectors of size n, or 2) matrices of size mxn, or 3) functions of a real variable, or even 4) differential operators form a vector space.

This is what we mean by abstraction. On the surface, all of the four spaces above may seem very different. Yet they have something deeply in common. Mathematicians love to find structures that seem different and yet have something in common, and then study the common structure all at once, rather than dealing with the unimportant surface differences.

We feel we are unfolding the very fabric that underlies everything.

Outline

1. Linear Combinations of Vectors
2. Linear Combinations of Matrices
3. Linear Combinations of Real Valued Functions
4. Linear Combinations of Operators
5. Vector Spaces
6. Elementwise Operations
7. Linear Systems of Equations

Linear Combinations by example

1. Linear Combinations of Vectors

If v_1 and v_2 are in $\mathbb{R}^n$ and c_1, c_2 are in $\mathbb{R}$ then $c_1v_1 + c_2v_2$ is a linear combination of v_1 & v_2.

Examples: (left n=5 with variables, right n=2 with numbers)

$$c_1\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_5 \end{bmatrix} + c_2\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_5 \end{bmatrix} = \begin{bmatrix} c_1x_1 + c_2y_1 \\ c_1x_2 + c_2y_2 \\ \vdots \\ c_1x_5 + c_5y_5 \end{bmatrix}$$

$$2\begin{bmatrix} 5 \\ 6 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 16 \\ 21 \end{bmatrix}$$

Remark: We'll see in a moment it's reasonable to talk about linear combinations of any number of vectors, not just two. (Notation: $\mathbb{R}$ denotes the real numbers.)

Linear Combinations by example

2. Matrices

Linear Combinations of Matrices

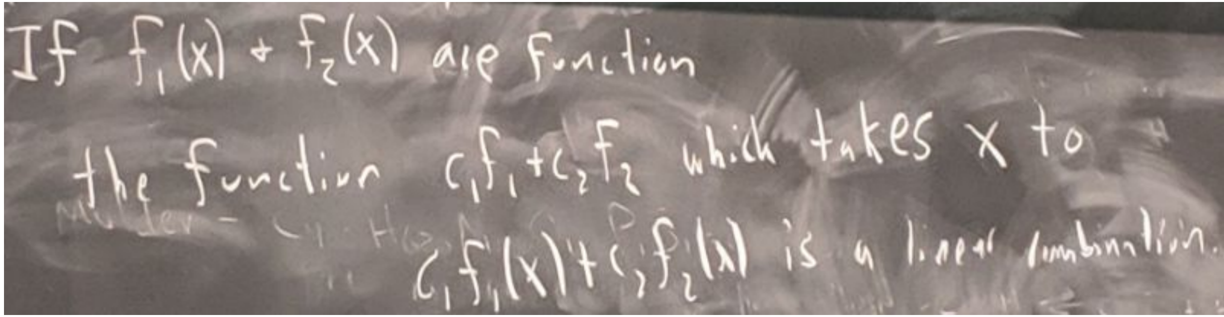
$A \in \mathbb{R}^{m,n}$ & $B \in \mathbb{R}^{m,n}$ $\quad c_1, c_2 \in \mathbb{R}$

$c_1 A + c_2 B$ is a linear combo. of A & B

$$2\begin{bmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{bmatrix} + 5\begin{bmatrix}0 & 1\\ 1 & 2\\ -1 & 0\end{bmatrix} = \begin{bmatrix}2 & 9\\ 11 & 18\\ -5 & 12\end{bmatrix}$$

← Silly me. Bottom Left entry should be +5.

Linear Combinations
By example



3. (Real valued) Functions

Consider functions like sin or pow_k defined as the function that takes x to sin(x) and x^k respectively.

Can we talk about 10*sin? Sure, it takes x to 10*sin(x). (the times “star” is optional)

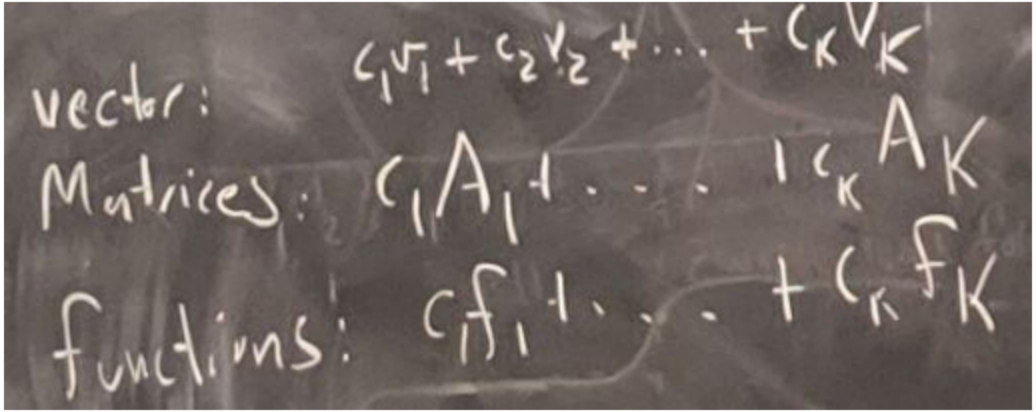
f = (10 * pow_2 + 3 * pow_4) ?

Sure we can, this function takes any x to $10x^2+3x^4$. We can write $f(x) = 10x^2+3x^4$.

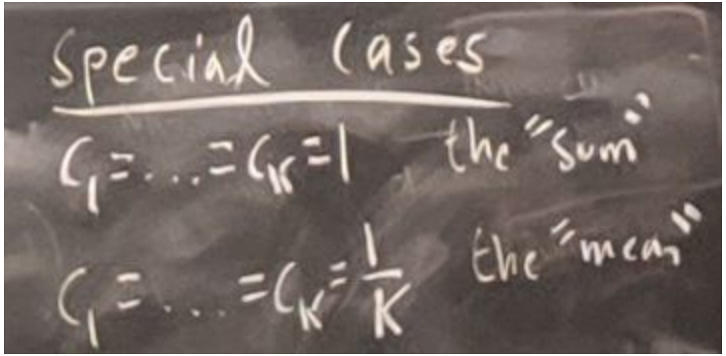
Thus (10 * pow_2 + 3 * pow_4)(2) = 10*4 + 3*16 = 88.

f = 5 exp + 2 log ? Sure $f(x)=5e^x+2\log(x)$.

Linear combinations can be any number of items



← vectors, matrices, and functions: all allow for linear combinations. Mathematicians: let's focus on what we have in common rather than our differences. (Editorial: All too often humans can take a lesson here.)

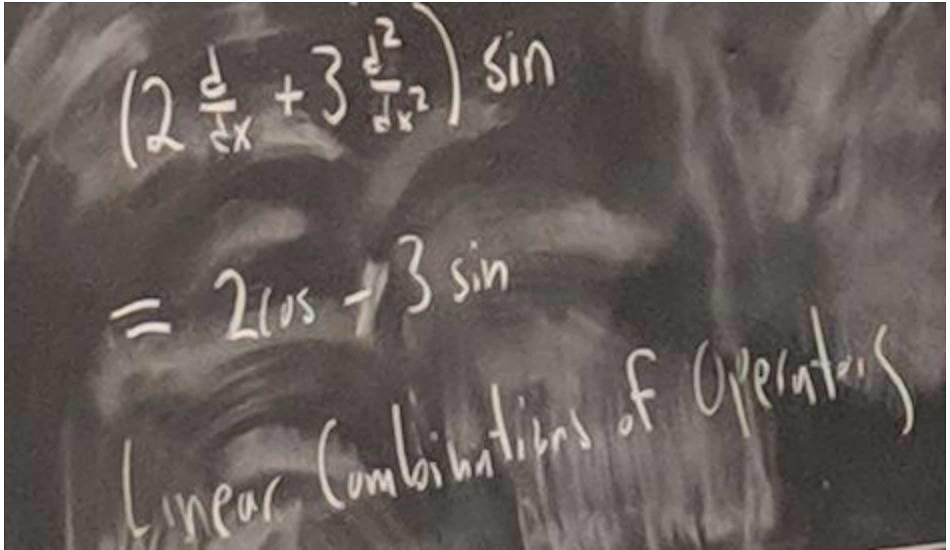


See VMLS page 18 for other examples.

Linear Combinations by example

Linear Combinations of Operators such as d/dx or d^2/dx^2.

$(2 * d/dx \ + 3 * d^2/dx^2)(f) = 2f' + 3f''$



Example when f = sin

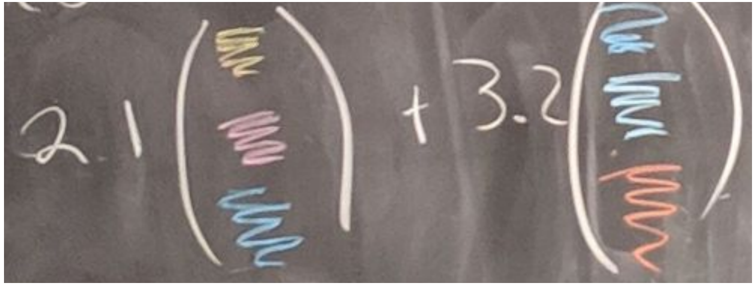
A function is a rule for taking numbers or vectors into new numbers or vectors.

An “operator” might be thought of as a “function function” in that it provides a rule for turning one function into another function.

Vector Spaces

Informal definition of a (real) vector space: Any mathematical set where it is sensible to take real linear combinations, and not go outside the set.

Not $\mathbb{Z}^3$ though, because real linear combinations of integers are not integers.



Also not colors[3] unless some meaning is given to quantities such as the linear combination on the left.

Elementwise Operations

Consider f(x) = sin(x), x^2, or e^x. What should we mean by f(vector) or f(matrix)??
Sometime one sees elementwise application.

For example

$$f\begin{pmatrix}1\\2\\3\end{pmatrix}=\begin{pmatrix}f(1)\\f(2)\\f(3)\end{pmatrix} \quad \text{or} \quad f\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}=\begin{pmatrix}f(1)&f(2)&f(3)\\f(4)&f(5)&f(6)\\f(7)&f(8)&f(9)\end{pmatrix}$$

The problem is that squaring a vector leads to confusion, e.g. is it the dot product, or the cross product (which really is a three dimensional thing only) or not really defined.
Squaring a matrix is even more confusing -- the matrix multiply or the elementwise square?

Also $e^A = I + A + A^2/2! + A^3/3! + A^4/4! + \ldots$.   Is a thing that is different from the elementwise exponential as every power of A denotes matrix multiply.

Elementwise notational solution

The simple little dot or point (“.”) which may be thought of as pointwise (synonymous with elementwise) nicely solves this problem.

$$f.\begin{pmatrix}1\\2\\3\end{pmatrix}=\begin{pmatrix}f(1)\\f(2)\\f(3)\end{pmatrix}$$

$$f.\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}=\begin{pmatrix}f(1)&f(2)&f(3)\\f(4)&f(5)&f(6)\\f(7)&f(8)&f(9)\end{pmatrix}$$

Elementwise notational solution

f.(vector) or f.(matrix) makes perfect sense elementwise for any function f .

Without the dot

if $f(x)=x^2$

f(vector) should be undefined, f(matrix) should be the matrix multiply square

And if $f(x)=e^x$

f(vector) should be undefined f(matrix) should be the matrix exponential.

Many authors rely on context to disambiguate. Benefits: saves the dots, Downside: can be confusing. Especially if you did want to write a software program. Some find the dots hard to read at first, but useful with practice.

Elementwise notational solution

Can be chained

Notation for a nonlinear neural network: h.(C* h.(B*h.(Ax)))

Example

h(x) is a nonlinear “activation function” such as h(x) = max(x,0). (“activates” when x is positive). (Old fashioned language that seems to be sticking, unfortunately: this h(x) is known as ReLU, or rectified linear unit function.)

Typical example x is an n_0-vector. A is an n_1 by n_0 matrix.

B is an n_2 by n_1 matrix, and C is an n_3 by n_2 matrix, etc.

Also elementwise

A ./ B is elementwise divide

A .+B is the same as A+B for ordinary vectors and matrices but can generalize

etc

One of the linear algebra “biggies”

What linear combination of vectors (meaning find the “c”’s) will produce a desired vector b?

This is equivalent to solving Ax=b, and has a row and column view.

Worth watching (probably at double speed) is Strang’s Lecture 1 Video.