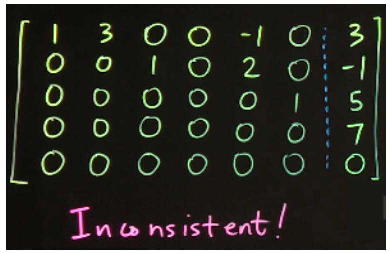
linear algebra (lin alg)

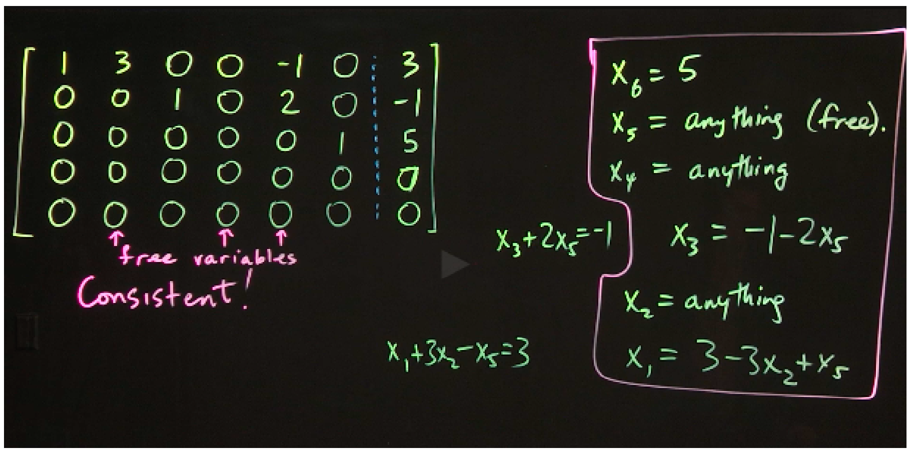
- lin alg deals with systems of linear equations
- we can write linear equations into augmented matrix form
- augmented matrix contains coefficients matrix on the left
- linear equations - left, augmented matrix - right

$$\begin{bmatrix} 10x_1 + 20x_2 = 50 \\ 3x_1 \; -7x_2 = 41 \end{bmatrix} \quad \left[\begin{array}{cc|c} 10 & 20 & 50 \\ 3 & -7 & 41 \end{array}\right]$$

- can do **3 types** of elementary row operations to solve linear equations
- types of row operations
    1. multiply row by a non-zero scalar or non-zero number
    2. add a scalar multiple of one row to another row
    3. swap two rows
- these row operations are reversible
- A and B are row-equivalent, if augmented matrix B is obtained from augmented matrix A via a set of elementary row operations
- If A and B are row-equivalent, the linear equations represented by A and B have the same solutions.
- when to stop doing row operations → when augmented matrix is in Reduced Row Echelon Form (RREF) i.e.
    1. In any non-zero row, the first non-zero entry is 1. (This is called a pivot.)
    2. Each pivot is the only non-zero entry in its column.
    3. A pivot in a lower row is to the right of any pivot in a higher row.
    4. All rows of all zeros (if any) are at the bottom.
- Gaussian elimination -- help put augmented matrix in RREF
    1. Swap two rows to get the left-most remaining non-zero entry to the topmost remaining row.
    2. Scale that row to make its leading entry a 1.
    3. Clear all non-zero entries in the column of that pivot (both above and below the pivot) by adding/subtracting multiples of that row to/from other rows.
    4. repeat the procedure with the remaining rows.
- RREF matrix interpretation
    - 0 0 0 0 …. 0 | non-zero → inconsistent (no solution)



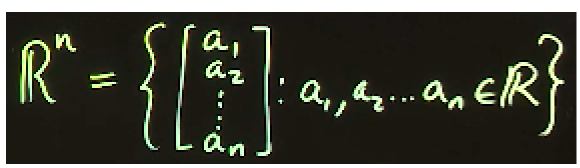
- if there are non pivot columns i.e. free columns or free variables → more than one solution



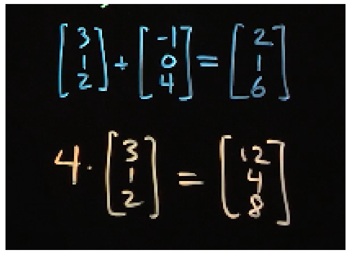
- If there are only pivot columns → only one (unique) solution for system of linear equations
- Maximum number of pivot columns is min(#row, #column)

vectors

- vector space R n can be defined as



- vector space = a set in which we can do two main operations and get member of the same set
  - addition -- vectors are added element wise
  - multiplication -- vectors elements multiplied by scalar

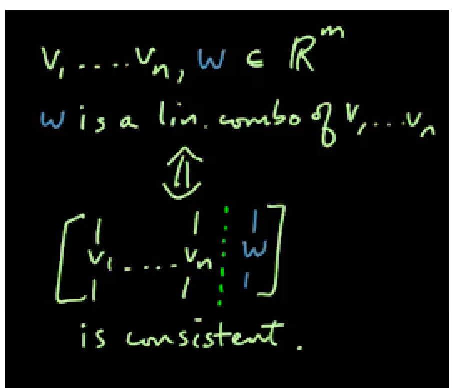


- vectors in different spaces (R n) are different e.g. [1,1,0,0] != [1,1,0]
- vectors are equal when they have same magnitude & direction regardless of their origin
- vector addition is **commutative** a+b = b+a and **associative** (a+b)+c = a+(b+c)
- scalar multiplication is **distributive** over vector addition a(u+v) = au+av
- vector multiplication is **distributive** over scalar addition (a+b)v = av+bv
- scalar multiplication is associative (ab)v = a(bv)
- scalar multiplication is like stretching or shrinking vectors
- vector multiply with negative scalar flips vector's direction

** u,v are vectors

linear combinations and spans

- if V is vector space over field F (some set e.g. R -- real, C -- complex), a linear combination of vectors v_1, v_2, …, v_n is $a_1v_1 + a_2v_2 + … + a_nv_n$, where the a_i are scalars in F .



- if augmented matrix of v_1, v_2, …, v_n | w in RREF has solution then w is linear combination of vs and $w_i = a_i$ where i = 1, 2, …, n
- set of all possible linear combinations of v_1, …, v_n is called the span of the vectors v_1, …, v_n
- span of list with no vectors (empty list) is zero vector
- v_1, …, v_n where v_i is member of R^m
  - if RREF of v_1, …, v_n has no rows with all zeros i.e. has pivot in every row v_1, …, v_n will span R^m i.e. span of v_1, …, v_n = R^m, every vector in R^m can be write as linear combination of v_1, …, v_n)

- when matrix column (n) < matrix row (m) -- $v_1, \ldots, v_n$ can not span R^m

matrix vector multiplication

- matrix times vector = linear combination of matrix columns with scalars multiple of columns being vector's elements

$$A = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix}, \quad x = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

$$Ax = a_1 v_1 + a_2 v_2 + \ldots + a_n v_n$$

Proposition. If $Ax = v$ is consistent, then the set of solutions to $Ax = v$ is a translation of the set of solutions to $Ax = \mathbf{0}$.

- $x = x_0 + y$ where $Ay = 0$ -- if Ax=v has at least a solution
- multiplication by matrix (A) distributes over vector addition A(v+w) = Av+Aw
- A(cx) = c(Ax) where c is a scalar and x is a vector

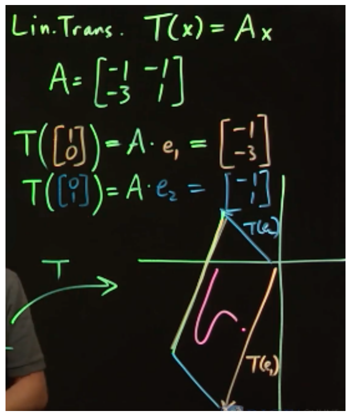
Linear Independence

- Ax = 0 has at least one solution and one of the solutions is x = zero vector (trivial solution)
- there can be solutions other than the trivial solution to Ax = 0 which are called non-trivial solutions
- vectors $v_1, \ldots, v_n$ are linearly independent if there exist only trivial solution to Ax = 0 where A's columns are $v_1, \ldots, v_n$
- vectors $v_1, \ldots, v_n$ are linearly dependent if there exist at least a non-trivial solution to Ax = 0 where A's columns are $v_1, \ldots, v_n$
- zero vector is linearly dependent -- zero times any number = 0 (non-trivial solution exist)
- if 2 vectors are collinear -- they are linearly dependent
- if v_1 is in span of $v_2, \ldots, v_n$ then $v_1, v_2, \ldots, v_n$ are linearly dependent
- if there exist at least one free variable (free column) in RREF of matrix A (many solutions for Ax = 0), then A's columns are linearly dependent
- matrix columns spans R^m if m < n where m = #rows, n=#columns i.e. Ax = b has at least one solution for any b in R^m

Linear Transformations

- takes a vector in domain space and give another vector in codomian space
  - T : V → W ; V = domain, W = codomain
- transformation can be written as a matrix
  - Ax = b where A is transformation matrix, x is input (member of domain space), b is output (member of codomain space)

- transformation is similar to function
- linear transformation (T) have the following properties
  - T(u+v) = Tu + Tv -- respects vector addition
  - T(av) = aT(v) -- respects scalar multiplication
- taking derivative is a linear transformation
- in lin alg linear transformation using matrix -- Ax = b is same as linear combination of matrix A columns with vector x
- standard matrix for transformation tells where the unit vector of each axis go after transformation (only applicable to vectors)
- identity transformation matrix (ones along main diagonal, zero elsewhere) keeps the vector unchanged after transformation.
- lines that start parallel always go in the same direction after linear transformation



Into and Onto

- for T : V → W, image of transformation T is set of all vectors that are member of codomain (W) and there exist at least a vector v in domain (V) which make T(v) = w
- T(v) = Av = w where A is standard transformation matrix of T
  - w is in image of T if Av = w is consistent (has at least one solution)
  - i.e. w is in span (set of possible linear combination) of columns of A
  - thus, image of T is span of A’s columns
  - span of matrix A’s columns = column space of A
  - if a column is linearly dependent on other columns then it doesn’t span (doesn’t add new information) column space
- T : V → W if image(T) = W then T is onto
- for T : $R^n \to R^m$ T is onto if columns of A spans R^m
- transformation is impossible to be onto when transformation matrix has #column < #row
- for T = Av if whenever T(u) = T(v), u is equal to v then T is into
- kernel of transformation T is set of v (vectors in domain) which make T(v) = 0
- if T is onto, then kernel(T) = zero vector (only one solution) because linear transformation of zero vector is zero vector itself
- if T is into then transformation matrix A of T has linearly independent columns
- transformation is impossible to be into when transformation matrix has #column > # row

Matrix Multiplication

- Composition of linear transformation is a linear transformation
- T o S (x) = T(S(x))
- if transformation T has transformation matrix A where A is m x n matrix (matrix with m rows, n columns), T is T : $R^n \rightarrow R^m$
    - #column = input dimension, #row = output dimension

Fundamental Fact of Matrix Multiplication. If the $m \times n$ matrix A is the standard matrix for $T : \mathbb{R}^n \to \mathbb{R}^m$, and the $n \times p$ matrix B is the standard matrix for $S : \mathbb{R}^p \to \mathbb{R}^n$, then AB will be an $m \times p$ matrix which is the standard matrix for $T \circ S : \mathbb{R}^p \to \mathbb{R}^m$.

- AB is defined when matrix A has #column = matrix B #row
- matrix multiplication is not commutative -- AB != BA
- matrix multiplication is associative -- A(BC) = (AB)C

Proposition. Suppose A is an $m \times n$ matrix and B is an $n \times p$ matrix. If the columns of A are linearly independent, and the columns of B are linearly independent, then the columns of AB are linearly independent.

Matrix Inverses

- inverse is used to undo transformation
- if A and B are inverses of each other, then AB = BA = identity matrix
- transformation is invertible when its transformation matrix A is invertible (A has inverse)
- inverse of A is written as A^{-1}
- zeros matrix is not invertible
- if transformation T is both into and onto, then the T is invertible
- inverse of a linear transformation is a linear transformation
- if matrix A's #row != #column, then A is not invertible
- though A is a square matrix, A can still not be invertible
- if a matrix A in RREF is an identity matrix, then A is invertible
- we can do RREF of [A | I] to get [I | A^{-1}]
- one sided inverse is not an inverse -- AB = I but BA != I
- if A and B are square matrices and AB = I, then BA = I and A and B are inverse of each other
    - AB = I -- B is right inverse of A
    - BA = I -- B is left inverse of A

More Matrix Operations

- if A and B are invertible, then AB is also invertible vice versa
- $(AB)^{-1} = B^{-1}A^{-1}$
- elementary matrix (E) is used to do single row operation on matrix (A), EA is matrix A that has been done a single row operation
- every elementary matrix is invertible (undoable) as row operation is reversible

- elementary matrices
    - scale a row (row i) by scalar non-zero c -- use identity matrix & replace identity matrix row i with row i * c
        - [1 0; 0 1] → scale row 2 by 3 → [1 0; 0 3]
    - add c x row i to row j -- use identity matrix and replace i th column of row j with c
        - [1 0; 0 1] → add 2 x row 1 to row 2 → [1 0; 2 1]
    - swap rows -- use identity matrix with rows swapped
        - [1 0; 0 1] → swap row 1 and row 2 → [0 1; 1 0]
- transpose of matrix A (A^T) is a matrix whose columns are the rows of A and whose rows are the columns of A i.e. A gets flipped across its main diagonal (from top left to bottom right)
- $(A^T)^T = A$
- $(AB)^T = B^TA^T$
- if A is invertible, then A^T is also invertible and vice versa
- A^T is into and onto if A is invertible
    - rows of A are linearly independent
    - rows of A span R^n if A is n x n matrix
- thus, for any n x n (square) matrix A to be invertible its rows must span R^n and must be linearly independent. This applies to A's columns as well.
- $(A^T)^{-1} = (A^{-1})^T$
- $(AB)^{-1} = B^{-1}A^{-1}$

Subspaces

- a subspace is subset of a vector space
- any vector space is subspace of itself
- criteria for W to be subspace of vector space V (let u,v be vector in W and c be any scalar)
    1. closed under addition -- u + v is still member of W
    2. closed under scalar multiplication -- cv is still member of W
    3. zero vector is member of W
- all 3 criteria above must be satisfied for a valid subspace
- set of only zero vector is a valid subspace
- vector spaces with different dimension can not be subspace of each other
- span of vectors in vector space V is a subspace of V
- every non-zero vector has at least 2 subspaces
    1. set of zero vector
    2. set of itself

Bases and Coordinates

- a list of vectors is basis of vector space V when the vectors are linearly independent and they span V
- unit vector $e_1, e_2, \ldots, e_n$ are basis (standard basis) of R^n
- vector space can have multiple set of basis vectors

- any basis of R^n must have exactly n vectors but any list of n vectors is not necessarily a basis of R^n
- to check if n vectors are the basis of R^n, put the vectors in a matrix (A) with each vector being A's column. If A is invertible then the vectors (columns of A) are one of the bases.
- if w is a vector in vector space V, there is exactly one way to write w as linear combination of V's basis vectors
- if w is linear combination of basis vectors B = $\{v_1, v_2, \ldots, v_n\}$ then w can be written as $w = a_1v_1 + a_2v_2 + \ldots + a_n v_n$ coordinate of vector w relative to basis vector $v_1, v_2, \ldots, v_n$ is a vector $[w]_B = [a_1 ; a_2 ; \ldots ; a_n]$
- $[v_1 \ v_2 \ \ldots \ v_n][a_1; a_2; \ldots; a_n] = w$ is same as writing w as linear combination of v_i where (i = 1,2 , …, n) and
  - v_i is other basis' vector of relative to current basis
  - a_i is vector in the other basis
  - w is vector of other basis relative to current basis

Dimension

- coordinatization
  - let B = $\{v_1, v_2, \ldots, v_n\}$ be basis vector for V and w is vector (linear combination of basis) in V
  - coordinatization = finding w relative to basis vectors in B (finding w in perspective of B)
- coordinatization C can be written as $C : V \to R^n$ -- $C(w) = [w]_B$
- coordinatization is a linear transformation which is onto and into
- every basis of a vector space has to have same number of vectors
- dimension = number of vectors in basis of vector space
- if vector space has no finite basis, then the vector space is infinite dimensional
- any vector space V has at most n linearly independent vectors where n = dimension(V)
- list of linearly independent vectors can be extended (remaining linear independence) until vector space is spanned to form basis if dimension of the vector space is finite
- for any W which is a subspace of vector space V where dimension(V) is **finite**, 0 <= dimension(W) <= dimension(V) and if dimension(W) = dimension(V) then W = V
- list of only zero vector {0} has dimension = 0
- we can remove some vectors that are linearly dependent on other vectors in list that spans a vector space to get basis of the vector space

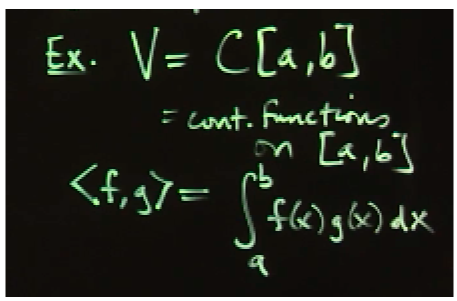
Rank and Nullity

- for a linear transformation $T : V \to W$ with standard matrix A
  - nullity of T = dimension of the kernel of T = dimension of the kernel of A
  - rank of T = dimension of the image of T = dimension of the image of A
  - dimension(V) = rank(T) + nullity(T) -- here dimension(V) is finite
- if nullity(T) = 0, then T is into (only trivial solution)
- If rank(T) = dimension(V) where V is vector space, then T is onto

- nullity(T) = #free variables in RREF of standard matrix (A) of transformation T
    - RREF [A | 0]
- Image(T) = span of A's columns, columns of A may contain basis of Image(T)
- rank(T) = #pivot columns in RREF of standard matrix (A) of transformation T
- the columns of A which correspond to the pivots in RREF of A will form a basis for the image of transformation T with standard matrix A
- dimension of domain = #columns of transformation matrix = #free columns + #pivot columns = #nullity + #rank

Inner Products

- dot product (inner product) of v and w = v.w = <v,w> = sum($v_i w_i$) where v,w are vectors in R^n and i = 1, 2, ..., n
- v.w = w.v
- cv.w = c(v.w) where c is a scalar
- v.v >= 0 always -- v.v = 0 only for zero vector
- (u+v).w = (u.w) + (v.w) where u,v,w are vectors in R^n
- dot product is used to measure how close vectors are (dot product is a scalar)
- length of vectors and distance between vectors are not preserved by most of linear transformations
- inner product of 2 continuous functions f(x) and g(x) over interval [a,b] is



    here we are treating function f, g as vectors
- magnitude or norm of vector v = ||v|| = sqrt(v.v) i.e. $||v||^2$ = v.v
- distance between 2 vectors v and w is ||v - w||
- if 2 vectors are orthogonal (perpendicular) then their inner product = 0
- if v and w are orthogonal vectors then from pythagoras theorem

$$\|v\|^2 + \|w\|^2 = \|v + w\|^2.$$

- list of vectors {v_1,v_2, …, v_n} is orthogonal if v_i.v_j = 0 where i != j i.e. v_i and v_j are not same vector
- list of vectors is orthonormal when the list is orthogonal and all vectors inside list have norm of 1
- any orthogonal list of **non-zero** vectors is linearly independent.
- every orthonormal list is linearly independent

Projections

- projection is like shadow of vector (v) onto subspace W
  - if v is inside subspace W then projection of V will be V it self
- having an **orthonormal** basis $\{w_1, w_2, \ldots, w_n\}$ for a subspace W, we can find projection vector v onto W as
  - $Proj_W(v)$ = sum(($v.w_i$) w_i) -- linear combination of W's basis with coefficients being shadow of v onto each basis (v dot w_i)
  - $Proj_W(v)$ is a vector in W that is an orthogonal projection of v
  - $Proj_W(v)$ - v is perpendicular to all vectors in W
  - there is only one vector in W which is $Proj_W(v)$
  - $Proj_W(v)$ is the most closest vector to v in W
- we can make orthogonal vectors to be orthonormal by dividing each vector with its norm to get vectors of length 1
- let $\{u_1, u_2, \ldots, u_k\}$ be list of **orthogonal** basis vectors

$$\mathrm{Proj}_W(v) = \sum_{i=1}^{k} \frac{\langle v, u_i \rangle}{\langle u_i, u_i \rangle} u_i .$$

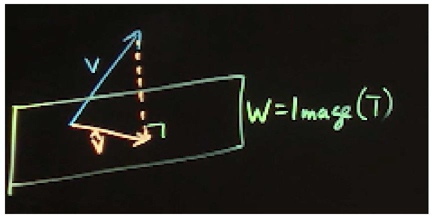
- if we know orthogonal non-zero basis vectors, then we can definitely calculate vector relative to those basis
- Gram-Schmidt Algorithm is used to find orthogonal basis from non-orthogonal basis vectors
  - if we have basis $\{w_1, w_2, \ldots, w_n\}$ our new orthogonal basis $\{u_1, u_2, \ldots, u_n\}$ will contain exactly n vectors where
    - $u_1 = w_1$
    - $u_i = w_i$ - Projection(w_i) on $\{u_1, \ldots, u_{i-1}\}$
- orthogonal projections can be done in any vector space with an inner product such as space of functions

Orthogonal Matrices

- $v.w = v^T w$ -- when v,w are treated as matrices with one column
- orthogonal matrix (A) have 4 properties (any of these is true then A is orthogonal)
  - A is an n x n matrix whose columns are orthonormal basis for R^n
  - $AA^T = A^TA = I$ which means $A^{-1} = A^T$
  - A preserves inner product (angle) of 2 vectors
    - $(Av).(Aw) = v.w$
  - A preserves norm (distance) of any vector i.e. for any vector v,w
    - $\| Av \| = \| v \|$

Least Squares

- least squares method find solution to $Ax = v$ i.e. find x that make Ax closest to v by minimizing distance from Ax to v
- one way to find x is to find projection of v onto subspace W constructed from linear combination A's column where A is standard matrix for transformation T



- $A^T \cdot (v - v_{proj}) = 0$ -- columns of A (rows of A^T) i.e. any vector in W are orthogonal to $v - v_{proj}$
- we get $A^Tv = A^Tv_{proj} = A^TAx$ -- here $v_{proj} = Ax$ as we are trying to find x to approximate v using v_{proj}