Midterm 1 Review

EE16A FALL 2018

Slides by Chris Mitchell and Nick Werblun
Modified by Vijay Govindarajan for Fall 2018

Outline

- Matrix Transformations and Gaussian Elimination
- Vector Spaces and Subspaces
- Linear (In)dependence, Span, and Basis
- Range, Column Space, and Null Space
- Eigenvalues and Eigenvectors
- Practice exam problems

Matrix Transformations and Gaussian Elimination

Matrix Transformations

- We can view Ax as a linear operation A applied to x
- Geometric operations
  - Rotation
  - Scaling
  - Reflecting
- Shifting/Translation?
  - Not a linear transformation
- Order of matrix multiplication matters (most of the time)
- See Dis 2B Worksheet and IPYNB for more

Gaussian Elimination

- I have a system of equations and I want to know the solution
    - Doing it by hand is hard though
- Solution: Build an augmented matrix and use Gaussian Elimination
- See Note 1 for more

$$\begin{bmatrix} 2x & + & 4y & + & 2z & = & 8 \\ x & + & y & + & z & = & 6 \\ x & - & y & - & z & = & 4 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 2 & 4 & 2 & 8 \\ 1 & 1 & 1 & 6 \\ 1 & -1 & -1 & 4 \end{array}\right]$$

Gaussian Elimination

- Goal: Reduced Form
  - When reduced, you can read the solutions right out of the matrix (if they exist)
- Apply basic row operations to get into this form or reach a stopping condition
- Pivot: Having a nonzero value in a column where all values left of it are 0.
  - In reduced form, all values above and below a pivot are also 0
  - NOTE: Not all matrices will have a pivot in every column

Pivot

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Pivot

NO Pivot

$x_1 = -2$

$x_2 = -4$

$x_3 = ???$

What??

Gaussian Elimination

Examples (once in reduced form)

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 0 \\ \boxed{0 \quad 0 \quad 0} & & & \boxed{0} \end{array}\right]$$

0=0, inf. solutions

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Pivot in every column, exact sol

$$\left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

0=1, no solutions

Vector Spaces and Subspaces

Vector Space (formally)

- Definition: A set of vectors V, scalars F, and vector addition and scalar multiplication operations with following properties (from Dis3B)

Exercise: Try to show these properties hold for V = R^n, F = R

- Vector Addition
  - Associative: $\vec{u}+(\vec{v}+\vec{w})=(\vec{u}+\vec{v})+\vec{w}$ for any $\vec{v},\vec{u},\vec{w}\in V$.
  - Commutative: $\vec{u}+\vec{v}=\vec{v}+\vec{u}$ for any $\vec{v},\vec{u}\in V$.
  - Additive Identity: There exists an additive identity $\vec{0}\in V$ such that $\vec{v}+\vec{0}=\vec{v}$ for any $\vec{v}\in V$.
  - Additive Inverse: For any $\vec{v}\in V$, there exists $-\vec{v}\in V$ such that $\vec{v}+(-\vec{v})=\vec{0}$. We call $-\vec{v}$ the additive inverse of $\vec{v}$.
- Scalar Multiplication
  - Associative: $\alpha(\beta\vec{v})=(\alpha\beta)\vec{v}$ for any $\vec{v}\in V$, $\alpha,\beta\in F$.
  - Multiplicative Identity: There exists $1\in F$ where $1\cdot\vec{v}=\vec{v}$ for any $\vec{v}\in F$. We call 1 the multiplicative identity.
  - Distributive in vector addition: $\alpha(\vec{u}+\vec{v})=\alpha\vec{u}+\alpha\vec{v}$ for any $\alpha\in F$ and $\vec{u},\vec{v}\in V$.
  - Distributive in scalar addition: $(\alpha+\beta)\vec{v}=\alpha\vec{v}+\beta\vec{v}$ for any $\alpha,\beta\in F$ and $\vec{v}\in V$.

Vector Space (informally)

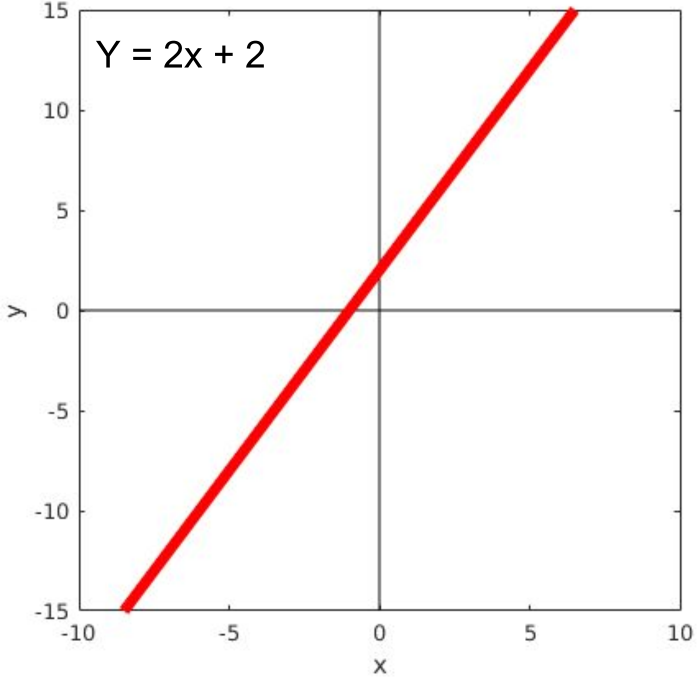
- A vector space V is a set of vectors that is closed under vector addition, scalar multiplication, and has a zero vector.
- This means if you take a linear combination of any vectors (v_1, v_2) in V, then the resulting vector (w) stays in V.

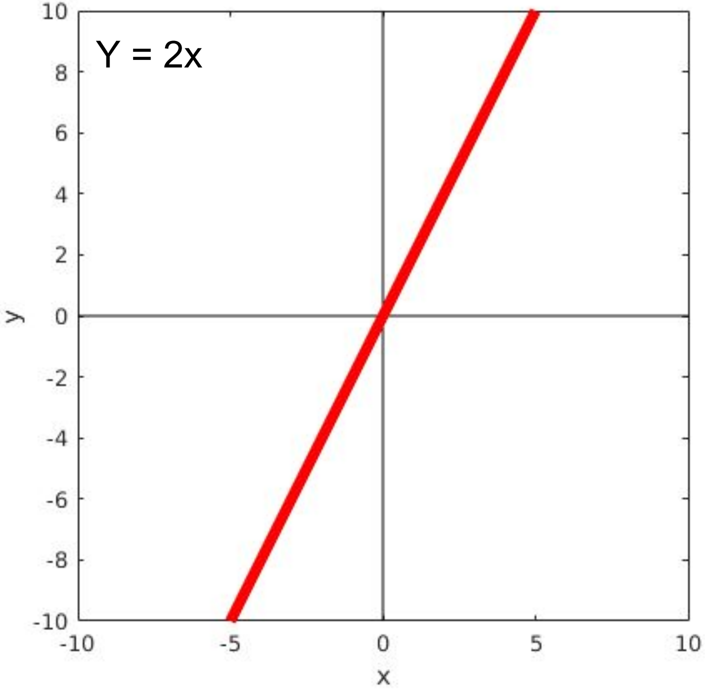
$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 = \vec{w}$$

Subspace

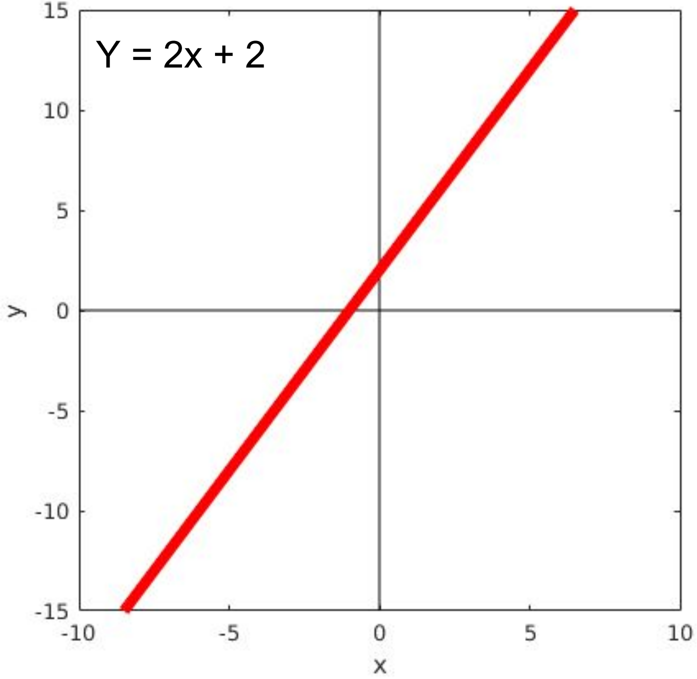
- A subset (W) of a vector space (V) is a **subspace** if:
    - W itself is a vector space, which we check with the following:
        - For any vectors u, v in W, and any scalar a:
            - u + v is also in W (closed under addition)
            - a*u is also in W (closed under scalar multiplication)
            - 0 is in W (existence of zero vector)
- **Note:** We refer to V as the parent vector space of the subspace W. Also, V is a subspace of itself.

Are these subspaces of V = R^2?

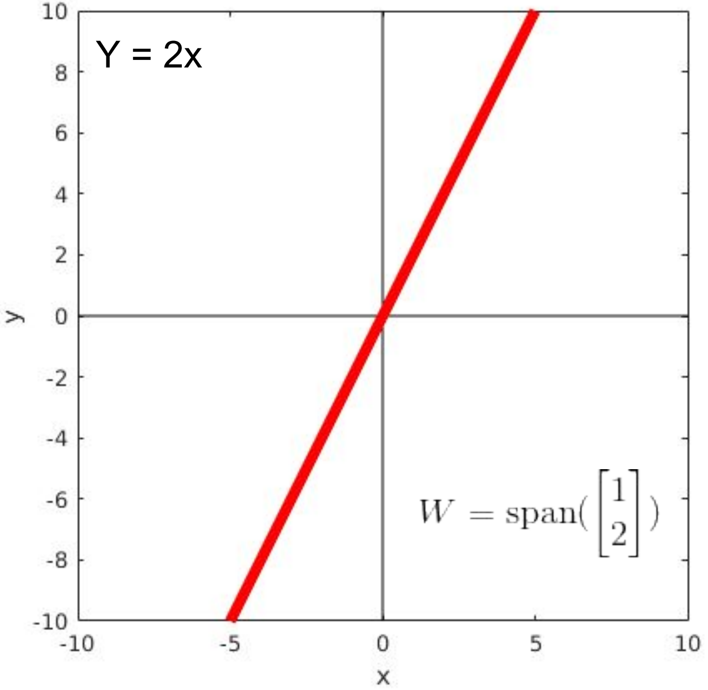




Are these subspaces of V = R^2?



No. Not closed under vector addition or scalar multiplication and doesn't contain the zero vector.



Yes. The line contains the zero vector. We can take linear combinations of points on the line and stay on the line.

Linear (In)dependence, Span, and Basis

Linear (In)dependence - Definition

- If a set of vectors is **linearly dependent**, that means one of them can be formed from a linear combination of the others (falls in the span of the other vectors) and is *redundant*.
- If a set of vectors is **linearly independent**, on the other hand, then no vector in that set can be constructed using linear combinations of the other ones.
- One way to determine this is to solve the equation: $\alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n = \vec{0}$

  If the only solution is alpha_1, ... , alpha_n = 0, then the set of vectors {v_1, …, v_n} is linearly independent.

$$\alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n = \vec{0} \Rightarrow \alpha_1, \cdots, \alpha_n = 0$$

Linear (In)dependence

Examples: Linearly dependent?

$$\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -4 \\ -2 \end{bmatrix}, \begin{bmatrix} 5 \\ -10 \\ 8 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Linear (In)dependence

Examples: Linearly dependent?

$$\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -4 \\ -2 \end{bmatrix}, \begin{bmatrix} 5 \\ -10 \\ -8 \end{bmatrix}$$

YES

$$-2\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 3 \\ -4 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 \\ -10 \\ -8 \end{bmatrix}$$

Thus the third vector is redundant information

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

NO. They're INDEPENDENT. The only way to solve

$$a_1\begin{bmatrix} 1 \\ -1 \end{bmatrix} + a_2\begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0$$

Is if BOTH coefficients are 0

Linear (In)dependence

NOTE: This boils down to essentially solving a system of equations. If we want to find the coefficients, we can do it by inspection, or:

$$a_1 \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 3 \\ -4 \\ -2 \end{bmatrix} + a_3 \begin{bmatrix} 5 \\ -10 \\ -8 \end{bmatrix} = \vec{0}$$

$$\left[\begin{array}{ccc|c} 2 & 3 & 5 & 0 \\ -1 & -4 & -10 & 0 \\ 1 & -2 & -8 & 0 \end{array}\right]$$

Span

- Span (Noun) = All linear combinations of a set of vectors {v1, …, vk}
    - Span({v1, .. vk}) = a1*v1 + … + ak*vk for any scalars a1, …, ak

- Span (Verb) / Spanning (Adj)
    - A list of vectors {v1,..,vk} spans a vector space, V, if every vector in V is in the span({v1,..,vk})

Basis and Dimension

- Basis: a set of vectors that are spanning + linearly independent
  - A minimal, spanning set of vectors for a given vector space
- Dimension: number of vectors in the basis
  - All possible bases for a vector space contain the same number of vectors
  - The number of vectors in the basis is the dimension
- Helpful fact: any set of n linearly independent vectors in a n-dimensional vector space is a basis for that space.

Range, Column Space, and Null Space

Range and Column Space

- Let A have columns A_1, …, A_n.
- Column Space(A) = Span({A_1, …, A_n})
- Range(A) = What you can “reach” with that matrix
    - AKA all y such that there exists an x where Ax = y
    - AKA Column space!
- These ideas are equivalent:
    - Range(A) = Column Space(A) = span({A_1, .., A_n})
- The dimension of Range(A) is called the rank of A.

Null Space

- Null(A) is the set of vectors that map to a zero output:
  - Null(A) = {x such that Ax = 0}
- This is related to the idea of linear independence:
  - If columns of A are linearly independent, then Null(A) = {0}, i.e. only the zero vector is in the null space
  - If columns of A are linearly dependent, then the nullspace is a subspace of dimension 1 or more

Why do Range(A) and Null(A) matter?

- One reason is that it helps us to understand solutions to $Ax = b$
- Existence of a solution
  - If b is not in the Range(A), then there is no solution to $Ax = b$
- Uniqueness of a solution
  - Suppose b is in the Range(A), so there is at least one solution.
  - Then Null(A) tells us if that solution is unique or not.
  - If Null(A) is just the zero vector, then we have a unique solution to $Ax = b$.
  - If Null(A) is a subspace of dimension 1 or more, we have infinite solutions to $Ax = b$
    - Why? Pick any non-zero vector in Null(A). Call it v_null.
    - Let $x = u$ be some solution to $Ax = b$.
    - $A(u + v\_null) = b + 0 = b$
    - So $x = u + v\_null$ is another solution to $Ax = b$. We can scale v_null and get an infinite set of solutions to $Ax = b$.

How to Find Null Space

- Begin with Gaussian Elimination for $Ax = 0$
- Look at constraints/equations in reduced form
- Identify free variables; write all other variables in terms of these
  - Usually, we select the variables that don't have corresponding pivots
- Write out the solution vector in terms of the free variables
- Factor out the free variables so you have a linear combination of vectors, with the free variables as the scalars
- Columns being scaled by free variables form a basis
- See [Dis 4A] for examples

Example: Null Space of A

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

Step 1: Gaussian Elimination

$$\mathbf{A}\vec{x} = \vec{0}$$

$$\begin{bmatrix} 1 & 4 & 7 & 0 \\ 2 & 5 & 8 & 0 \\ 3 & 6 & 9 & 0 \end{bmatrix}$$

Step 2: Identify free variables (z in this case)

$$\begin{bmatrix} 1 & 4 & 7 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Step 3: Write solution in terms of free parameters

$$z = t, t \in \mathbb{R}$$
$$y = -2t$$
$$x = -7z - 4y = t$$

Step 4: Factor free parameters

$$\vec{x} = \begin{bmatrix} t \\ -2t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

Step 5: Find a basis for N(A)

$$N(A) = \text{span}\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \}$$

Is it Invertible?

- Matrix or vector division does not exist!!!
- Only square matrices can have inverses (if they are invertible).
- **A** is not invertible if it
  - has a non-trivial null space (a null space that doesn’t just contain zero)
  - has a zero eigenvalue (more on eigenvalues later)
  - is not full rank (rank of matrix not equal to dimension)
  - has determinant zero (more on determinants later)
- If any of the statements in the above list are false, then **A** is invertible.

Determinants

- For this class, you need to know 2x2 determinants

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc$$

- Determinants encode some very neat info about **square** matrices
- They are linked to the area (2D) of the parallelogram or volume (3D) or parallelliped formed by the rows of the A matrix
- Important for finding eigenvalues of a matrix A

Eigenvalues and Eigenvectors

Eigenvectors/Eigenvalues

$$A\vec{v} = \lambda\vec{v}$$

Eigenvectors/Eigenvalues

- An eigenvector of matrix A is a vector such that applying matrix A to it yields the SAME exact vector, but scaled by a constant.
- The constant that the vector is scaled by is the ***eigenvalue***

Eigenvectors/Eigenvalues

- Find the eigenvalues by solving the equation resulting from:

$$\det(A - \lambda I) = 0$$

- You will get some polynomial whose roots are the eigenvalues
- Some notes:
    - The 0 vector can never be an eigen**vector**
    - You can definitely have repeated eigen**values**

Eigenvectors/Eigenvalues

- To find the eigenvectors, you first must find the eigenvalues
- Once obtained, plug in the eigenvalues (one at a time) into

$$(\mathbf{A} - \lambda I)\vec{v} = \vec{0}$$

- Note that this is now finding the null space of the resulting matrix $\mathbf{A} - \lambda I$

Imaging Lab (Basically More Eigenvectors)

- We have some sensor readings, s, related to our image by:

$$\vec{s} = H\vec{\imath} + \vec{\omega}$$

Noise. Random. Unknown. Hated

Sensor readings

Matrix of image masks

Image

- Without noise, we solve by multiplying by H inverse on both sides to get i
- With the noise, we have to care about the product of H inverse and w

Imaging Lab (Basically More Eigenvectors)

- For the NxN “mystery” matrix we used in imaging 3, it is constructed by design to have N distinct eigenvalues
    - This implies that there are N distinct eigenvectors
- With N independent vectors, we can span N dimensional space
- This implies that the noise (which is an N length vector) can be spanned by our eigenvectors

Imaging Lab (Basically More Eigenvectors)

- Thus we can say,

$$\vec{\omega} = \alpha_1 v_1 + \alpha_2 \vec{v_2} + \cdots + \alpha_n \vec{v_n}$$

- If we then apply H inverse when recovering our image, we get,

$$H^{-1}\vec{\omega} = \frac{1}{\lambda_1}\alpha_1\vec{v_1} + \frac{1}{\lambda_2}\alpha_2\vec{v_2} + \cdots + \frac{1}{\lambda_n}\alpha_n\vec{v_n}$$

- Since we know that if $A \rightarrow \lambda$ then $A^{-1} \rightarrow \frac{1}{\lambda}$ we know that the eigenvalues of our masking matrix matter.
    - Large eigenvalues of H leads to noise being lowered overall
    - Small eigenvalues of H leads to noise being amplified

Diagonalization

- Useful in computation. Matrix multiplication is computationally expensive, diagonalized matrices are easy to raise to powers.
- An NxN matrix is diagonalizable if it has n linearly independent eigenvectors

Diagonalization

- To find the diagonal form, you first need to find the eigenvalues and their corresponding eigenvectors.
- A diagonalized matrix is of the form $A = PDP^{-1}$
- P is of the form: $\begin{bmatrix} | & \cdots & | \\ \vec{a}_1 & \cdots & \vec{a}_n \\ | & \cdots & | \end{bmatrix}$
  - Each a vector is an eigenvector of A
- D is of the form: $\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$
  - The eigenvalues must be in the same relative order as you place the eigenvectors in P
  - Every other element is 0

Diagonalization

- Why is it useful?
- To compute A to a high power...

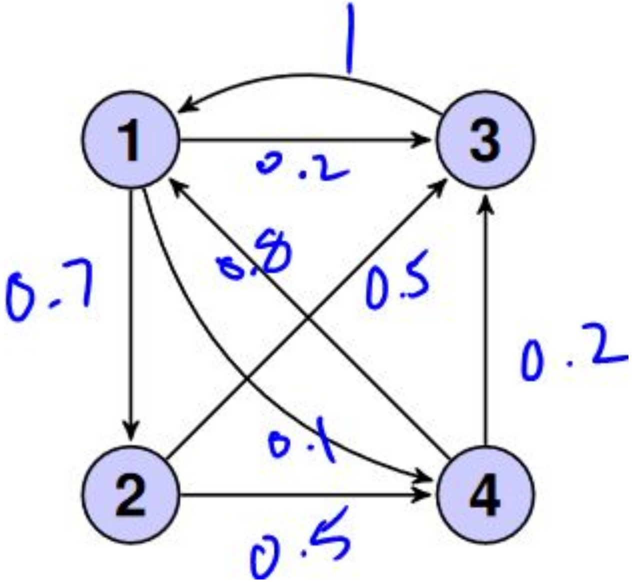
$$A^r = PD^rP^{-1}$$

$$A^r = P\begin{bmatrix} \lambda_1^r & & \\ & \ddots & \\ & & \lambda_n^r \end{bmatrix}P^{-1}$$

Flow Matrices and Graphs

- A system that updates its state every “time step”
- Usually represented by nodes connected by directed edges with some “weight”
    - Weight can mean different things depending on the system. Water tanks -> percentage of water that goes to a different (or the same) tank
    - In the pagerank problem, all edges leaving a node are weighted equally. Therefore if node A has 3 edges leaving, then each edge has weight ⅓
- Don’t forget to normalize page rank “importance scores”

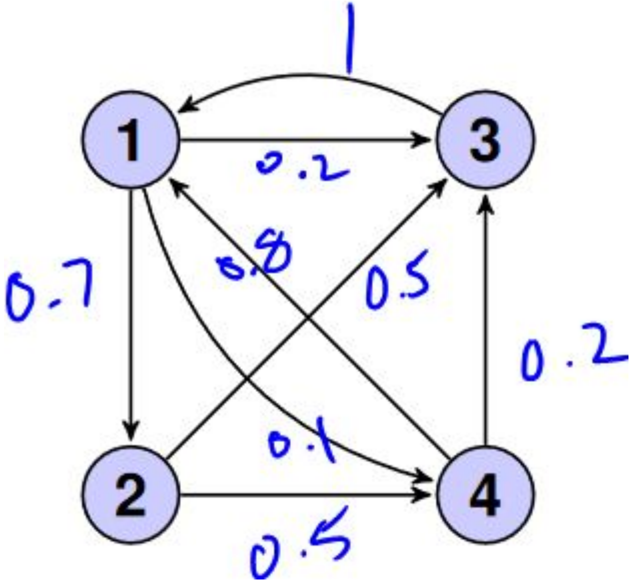
Flow Matrices and Graphs



Node 1 to all other nodes

$$T = \begin{bmatrix} \; \end{bmatrix}$$

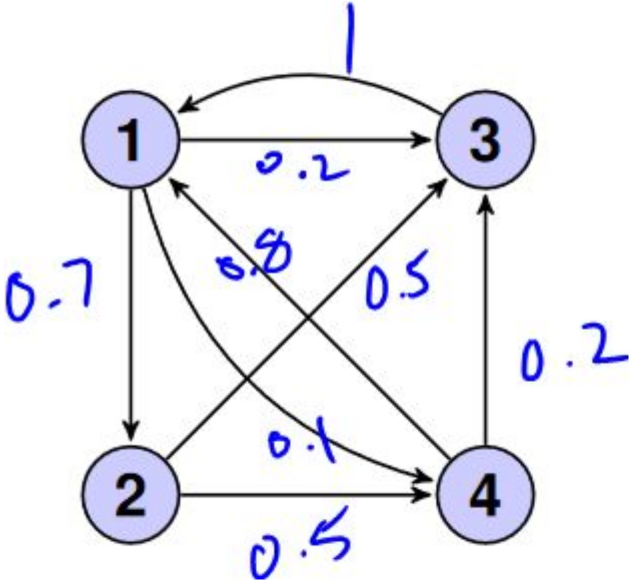
Flow Matrices and Graphs



$$T = \begin{bmatrix} \quad & & & \quad \end{bmatrix}$$

All other nodes to node 1

Flow Matrices and Graphs



Node 1 to all other nodes

$$T = \begin{bmatrix} 0 & 0 & 1 & 0.8 \\ 0.7 & 0 & 0 & 0 \\ 0.2 & 0.5 & 0 & 0.2 \\ 0.1 & 0.5 & 0 & 0 \end{bmatrix}$$

All other nodes to node 1

Flow Matrices and Graphs

- Conservative system? Note that these are COMMON but not GUARANTEED
  - No “stuff” leaves as time progresses. A system is conservative if the columns sum to 1
  - A column with sum >1 has “stuff” entering from outside. NOT conserved, states go to infinity
  - A column with sum < 1 has “stuff” leaking out. States will trend towards 0.
- Steady state?
  - We have a state as a vector s[n]
  - Transition matrix T describes how the state changes at each time step.

Flow Matrices and Graphs

- Next states? Previous states?
  - We have a state as a vector s[n] which represents how much stuff is at which node
  - Transition matrix T describes how the state changes at each time step.

$$T\vec{s}[n] = \vec{s}[n+1]$$

  - We can advance many timesteps by computing powers of T

$$T^N\vec{s}[n] = \vec{s}[n+N]$$

  - We can go backward in time to recover previous states IF T IS INVERTIBLE

$$T^{-1}\vec{s}[n+1] = \vec{s}[n]$$

Flow Matrices and Graphs

- Steady state? For pagerank it's called "importance score"
  - We're interested in a state where updating to the next step does not change the state vector
  - In other words…

$$T\vec{s}[n] = \vec{s}[n+1] = \vec{s}[n]$$

  - Maybe it looks more familiar written this way...

$$A\vec{v} = \vec{v}$$

Flow Matrices and Graphs

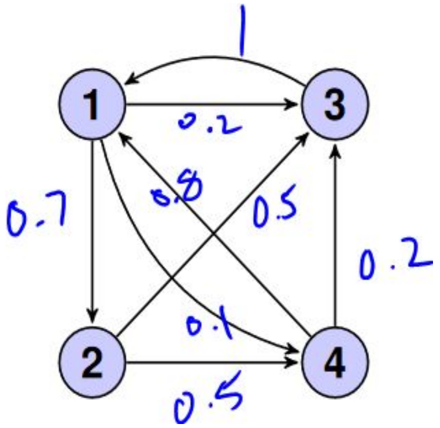
- Steady state?
  - We're interested in a state w…
  - In other words…

$$A\vec{v} = \lambda\vec{v}$$

Flow Matrices and Graphs

- We want to find the vector that does not change when T operates on it.
- We can find it with a foolproof (not fun) method of taking A to the nth power times s[n], then take the limit as n approaches infinity
- OR we can find the steady state with linear algebra.
    - We want the eigenvector that has an eigenvalue of 1. This vector will not change with T.
- NOTE: An eigenvalue of 1 does not mean there is always a steady state
    - we also want the other eigenvalues to be less than 1 in magnitude so that as n goes to infinity, λ^n goes to 0.
    - Diagonalize the matrix to see this in action, or represent the state vector as a combo of eigenvectors

Flow Matrices and Graphs



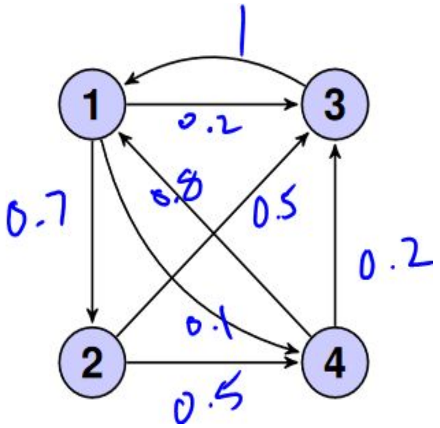
$$T = \begin{bmatrix} 0 & 0 & 1 & 0.8 \\ 0.7 & 0 & 0 & 0 \\ 0.2 & 0.5 & 0 & 0.2 \\ 0.1 & 0.5 & 0 & 0 \end{bmatrix}$$

- There's an eigenvalue of 1, and a few others with magnitude less than one. Thus a steady state exists
- The corresponding eigenvector is:

$$\vec{v_1} = \begin{bmatrix} 0.6897 \\ 0.482 \\ 0.441 \\ 0.3103 \end{bmatrix}$$

- There's just one last thing to do...

Flow Matrices and Graphs



$$T = \begin{bmatrix} 0 & 0 & 1 & 0.8 \\ 0.7 & 0 & 0 & 0 \\ 0.2 & 0.5 & 0 & 0.2 \\ 0.1 & 0.5 & 0 & 0 \end{bmatrix}$$

$$\overrightarrow{v_1} = \begin{bmatrix} 0.6897 \\ 0.482 \\ 0.441 \\ 0.3103 \end{bmatrix} \rightarrow \begin{bmatrix} 0.3584 \\ 0.2508 \\ 0.2293 \\ 0.1612 \end{bmatrix}$$

- These numbers don't make sense for our system. We need to "normalize" it so that the total sums to 1, representing 100%
- Divide each entry by the sum of all values

- Therefore at steady state, 35.84% of the stuff is in node 1, 25.08% in node 2, etc.

Derivations and Proofs - Tips and Tricks

- Proofs are generally not expected to be “hardcore”
    - Past exams also usually have “direct” proofs rather than something like contradiction, etc.
- Oftentimes you can do it in a few lines with some explanation
- If your proof becomes a bit too unwieldy, try to start over with a different strategy
- Start with the fundamental equations about what you know. Try to slowly involve the things you care about showing

Derivations and Proofs - Example

Show If matrix A has eigenvalues λ_i then A^{-1} has eigenvalues $\frac{1}{\lambda_i}$

Start with what you know: $A\vec{v} = \lambda\vec{v}$

$\rightarrow A^{-1}A\vec{v} = A^{-1}\lambda\vec{v}$

$\rightarrow I\vec{v} = A^{-1}\lambda\vec{v}$

$\rightarrow A^{-1}\lambda\vec{v} = \vec{v}$

$\rightarrow \boxed{A^{-1}\vec{v} = \frac{1}{\lambda_i}\vec{v}}$

General Tips

There are many linear algebra theorems and shortcuts. Don’t be afraid to Google some linear algebra cheat sheets!

But also be aware that you can only assume things that were shown on homework and in class. You cannot reference some theorem that was not covered as your proof.

Try to spend your time answering things you know well, and feel free to skip parts if you get stuck. You can always go back to the tougher parts if you have time.

Good luck!

Practice Exam Problems

Past Midterms

- SUMMER 2017 Q8 and Q9
  - Exam: https://d1b10bmlvqabco.cloudfront.net/attach/j6h2cdky5zk7c6/ib2kjoxfjiu1u1/j83u60vcykfx/mt1_531_cory.pdf
  - Solution: https://d1b10bmlvqabco.cloudfront.net/attach/j6h2cdky5zk7c6/ib2kjoxfjiu1u1/j83u7h6f3vwz/solmt1_1.pdf
- FALL 2017 Q7, Q8, and Q9
  - Exam: http://inst.eecs.berkeley.edu/~ee16a/fa17/homework/mt1.pdf
  - Solution: https://d1b10bmlvqabco.cloudfront.net/attach/j6h2cdky5zk7c6/ib2kjoxfjiu1u1/j8lycevulknm/solmt1_3.pdf
- If there is time, we can do more