Week 3

???

3.1 Linear Algebra Review

10/10:

- Today: Review of linear algebra.
- Start with a **vector space** over $\mathbb{R}$ or $\mathbb{C}$ or, more generally, any field K.
- **Vector space** (over K): A set equipped with addition and scalar multiplication such that the following axioms are satisfied.
  1. Commutativity and associativity of addition.
  2. Additive identity and inverse.
  3. Compatibility of scalar multiplication and addition (distributive laws).
  4. The additive identity times any vector is zero.
- In $\mathbb{R}^n, \mathbb{C}^n$, addition is component-wise and scalar multiplication is scaling of the element.
- For a homogeneous equation

$$y' = A(t)y = \begin{pmatrix} a_{11}(t)y^1 + a_{12}(t)y^2 + \cdots \\ \vdots \end{pmatrix} \begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{pmatrix}$$

  - If y_1, y_2 are solutions, any linear combination of them is a solution. This is called the **solution space** of the equation.
- **Linearly independent** (set of vectors): A set of vectors $x_1, \dots, x_m \in V$ such that the only coefficients $\lambda_1, \dots, \lambda_m$ such that

$$\lambda_1 x_1 + \cdots + \lambda_m x_m = 0$$

  is $\lambda_1 = \cdots = \lambda_m = 0$.
  - $\lambda_m \neq 0$ implies

$$x_m = -\frac{1}{\lambda_m}(\lambda_1 x_1 + \cdots + \lambda_{m-1} x_{m-1})$$

- **Maximal linear independence group**: A subset $X \subset V$ such that for any $y \in V$, $\{y\} \cup X$ is not linearly independent. *Also known as* **basis**.
- Theorem: Any basis in V has the same cardinality.
- **Dimension** (of V): The cardinality given by the above theorem. *Denoted by* $\mathbf{dim}\, \boldsymbol{V}$.

- We usually denoted a basis as an ordered n-tuple since the order often matters (for orientation?).
- Notational onventions.
  - For $\mathbb{R}^n, \mathbb{C}^n$, we will always use column vectors.
  - $x_1, x_2, \dots$ denotes vectors.
  - $x^1, x^2, \dots$ denotes the components of a column vector.
  - A vector component squared may be denoted $(x^1)^2$.
- **Standard basis** (for $\mathbb{R}^n$): The set of n vectors of length n which have a 1 as one entry and a zero in all the others and are all distinct.
- **Linear transformation** (of V to V): A mapping $\phi : V \to V$ satisfying

$$\phi(\lambda x + \mu y) = \lambda\phi(x) + \mu\phi(y)$$

- A mapping is completely determined by its action on the basis vectors:

$$\phi\left(\sum_{k=1}^{n} x^k e_k\right) = \sum_{k=1}^{n} x^k \phi(e_k)$$

- **Matrix** (of a linear transformation wrt. the standard basis): The $n \times n$ array

$$\begin{pmatrix} \phi(e_1) & \cdots & \phi(e_n) \end{pmatrix}$$

- If $\phi, \psi : V \to V$ are linear, $\phi \circ \psi$ is also linear.
  - Composition of linear transformations corresponds to matrix multiplication.
- Matrix multiplication: If

$$B = \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix}$$

  then

$$AB = \begin{pmatrix} Ab_1 & \cdots & Ab_n \end{pmatrix}$$

  where

$$Ax = \begin{pmatrix} a_{11}x^1 + \cdots + a_{1n}x^n \\ \vdots \\ a_{n1}x^1 + \cdots + a_{nn}x^n \end{pmatrix}$$

- We can talk about matrix inverses: If it exists, it is unique, and

$$AA^{-1} = A^{-1}A = I_n$$

- Matrix multiplication is not commutative in general. Shao gives a counterexample.
- A is invertible iff the columns of A are a basis for $\mathbb{R}^n$ (resp. $\mathbb{C}^n$).
- **Determinant** (of A): Not defined.
- Properties of the determinant.
  - Multilinear.

$$\det\begin{pmatrix} a_1 & \cdots & \lambda a_k + \mu\tilde{a}_k & \cdots & a_n \end{pmatrix} = \lambda\det\begin{pmatrix} a_1 & \cdots & a_k & \cdots & a_n \end{pmatrix} + \mu\det\begin{pmatrix} a_1 & \cdots & \tilde{a}_k & \cdots & a_n \end{pmatrix}$$

  - Skew-symmetric.

$$\det\begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_j & \cdots & a_n \end{pmatrix} = -\det\begin{pmatrix} a_1 & \cdots & a_j & \cdots & a_i & \cdots & a_n \end{pmatrix}$$

- Theorem: The determinant is uniquely characterized by these three (??) axioms.
- $\det I_n = 1$.
- Shao goes over computing the determinant via minors.
- Special cases:
  - If the matrix is upper- or lower-triangular, the determinant is equal to the product of the diagonal entries.
  - If the matrix is blocked upper- or lower-triangular, e.g.,

$$A = \begin{pmatrix} A_1 & * \\ 0 & A_2 \end{pmatrix}$$

    then $\det A = \det A_1 \cdot \det A_2$.
- $\det(AB) = \det(A)\det(B)$.
- $\det A \neq 0$ iff A is invertible.
- Direct formula to compute the inverse.

$$A^{-1} = \frac{1}{\det A}\left(a_{\ell k}(-1)^{k+\ell}\det A_{k\ell}\right)$$

  - Tedious for higher-dimensional cases, but quite sufficient for $n = 2, 3$.
- Let A be $n \times n$, and let $Ax = b$.
  - If A is invertible, then $x = A^{-1}b$.
  - If A is not invertible and $b \in \operatorname{span}(a_1, \dots, a_n)$, then $x = x_h + x_p$ where $Ax_h = 0$ and $Ax_p = b$.
- **Kernel** (of A): The set of all vectors $y \in \mathbb{R}^n$ (resp. $\mathbb{C}^n$) such that $Ay = 0$.
- **Range** (of A): The set of all linear combinations of $a_1, \dots, a_n$.
- Suppose $\phi : \mathbb{R}^n \to \mathbb{R}^n$ has matrix A under $(e_1, \dots, e_n)$. Let $(q_1, \dots, q_n)$ be another basis.
  - There exists a matrix Q such that $q_k = Qe_k$. Q is called the **connecting matrix** between $(e_1, \dots, e_n)$ and $(q_1, \dots, q_n)$.
  - Claim: Let $x \in \mathbb{R}^n$ have representation $x = (x^1, \dots, x^n)$ under the standard basis. Then under the Q basis, x has representation $x' = Q^{-1}(x^1, \dots, x^n)$. Similarly, $x = Qx'$.
  - Claim: ϕ has matrix $B = Q^{-1}AQ$ with respect to the Q basis.
- Matrix similarity: $A \sim B$ iff there exists Q invertible such that $B = Q^{-1}AQ$.
  - Implies that A and B describe the same matrix under different bases.
  - Matrix product under the old and new bases are related.

$$Q^{-1}ABQ = (Q^{-1}AQ)(Q^{-1}BQ)$$

  - Similarity preserves the determinant:

$$\det\left(Q^{-1}AQ\right) = \det\left(Q^{-1}\right)\det(A)\det(Q) = \det(A)\det\left(Q^{-1}\right)\det(Q) = \det(A)$$

3.2 Diagonalization and Jordan Normal Form

10/12:
- Similar matrices and Jordan Normal Form (JNF).
- Suppose $A : \mathbb{C}^n \to \mathbb{C}^n$ is linear. We can express A in a different basis with the help of the connecting matrix Q.
- In this lecture, we seek to find the most convenient basis in which to discuss our linear transformation.
- Today we will work in $\mathbb{C}^n$ (but all results hold for $\mathbb{R}^n$, too).
- **Invariant subspace** (of A): A subspace $K \subset \mathbb{C}^n$ such that $A(K) = K$.
- Suppose you have m invariant subspaces $K_1, \dots, K_m \subset \mathbb{C}^n$ whose pairwise intersection is $\{0\}$.
- **Direct sum** (of $K_1, \dots, K_m$): The collection of all vectors which can be represented as sums from each of the subspaces. *Denoted by* $\boldsymbol{K_1 \oplus \cdots \oplus K_m}$. *Given by*

$$K_1 \oplus \cdots \oplus K_m = \left\{ x \in \mathbb{C}^n \mid x = \sum_{j=1}^{m} x_j,\ x_j \in K_j \right\}$$

- Suppose $K_1, K_2 \in \mathbb{C}^n$ are invariant subspaces of A of dimension n_1, n_2, respectively, such that $K_1 \oplus K_2 = \mathbb{C}^n$. Then choosing a basis for K_1 and K_2, the matrix A takes the form

$$\begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}$$

  where B_1 is an $n_1 \times n_1$ block and B_2 is an $n_2 \times n_2$ block.
- **Eigenvalue** (of A): A complex number $\lambda \in \mathbb{C}$ such that $A - \lambda I$ is not invertible. *Denoted by* $\boldsymbol{\lambda}$.
  - Equivalently, $\det(A - \lambda I) = 0$.
- **Characteristic polynomial**: The polynomial in z defined as follows. *Denoted by* $\boldsymbol{\chi_A(z)}$. *Given by*

$$\chi_A(z) = \det(A - zI)$$

  - Similar matrices have the same characteristic polynomials.
- **Spectrum** (of A): The set of all eigenvalues of A.
- **Eigenvector** (of A): A vector $v \in \mathbb{C}^n$ corresponding to an eigenvalue λ via

$$Av = \lambda v$$

- Claim: The set of all eigenvectors corresponding to λ form an invariant subspace.

  Proof.
$$A(v_1 + v_2) = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2)$$
  □

- **Eigenspace** (of λ): The vector subspace of $\mathbb{C}^n$ equal to the span of the eigenvectors of λ. *Denoted by* $\boldsymbol{V_\lambda}$.
- **Algebraic multiplicity** (of λ): The degree of the $(z-\lambda)$ term in the factorization of the characteristic polynomial. *Denoted by* $\boldsymbol{\alpha_\lambda}$.
- **Geometric multiplicity** (of λ): The dimension of the eigenspace of λ. *Denoted by* $\boldsymbol{\gamma_\lambda}$.

- $\gamma_\lambda \leq \alpha_\lambda$.
- If $\alpha_\lambda = \gamma_\lambda$ for each λ, then each eigenspace V_λ has a basis such that $\oplus_\lambda V_\lambda = \mathbb{C}^n$.
  - Under this basis, the matrix of A is diagonal with all λ's (along the diagonal) repeated according to their algebraic multiplicity.
- **Superdiagonal**: The set of entries in a matrix which are directly above a diagonal entry.
- **Jordan block**: A $d \times d$ matrix corresponding to an eigenvalue λ that has λ as every diagonal entry, 1 as every superdiagonal entry, and zeroes everywhere else. *Denoted by* $\boldsymbol{J_d(\lambda)}$.
  - The geometric multiplicity γ_j is the number of Jordan blocks with eigenvalue λ_j. Of course, when $\gamma_j = \alpha_j$ (in particular, if $\alpha_j = 1$), there is no Jordan block corresponding to λ_j at all.
- We have that

$$
\begin{aligned}
J_d(\lambda)\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} &= \begin{pmatrix}\lambda\\0\\\vdots\\0\end{pmatrix} = \lambda\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} = \lambda e_1\\
J_d(\lambda)\begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix} &= \begin{pmatrix}1\\\lambda\\0\\\vdots\\0\end{pmatrix} = e_1 + \lambda e_2\\
&\vdots\\
J_d(\lambda)e_{d-1} &= e_{d-2} + \lambda e_{d-1}
\end{aligned}
$$

- For any linear transformation, we can find a basis such that the matrix is the diagonalized Jordon blocks.
- Theorem: Let A be an $n \times n$ complex matrix. Then there is a **Jordan basis** Q under which

$$
Q^{-1}AQ = \begin{pmatrix} J_{h_1}(\lambda_1) & & \\ & J_{h_{d_1}}(\lambda_1) & \\ & & \ddots \end{pmatrix}
$$

  - We have that $h_1 + \cdots + h_{d_1} = \alpha_1$??
  - The proof will not be tested — it is very hard. Shao will sketch it, though.
  - The proof is constructive: It will tell you how to convert a matrix into the Jordan normal form.
- Proof procedure:
  1. Determine the eigenvalues as well as their algebraic and geometric multiplicities.
     (a) Compute $\chi_A(z)$.
     (b) Find $\lambda_1, \dots, \lambda_m$ (factor $\chi_A(z)$).
     (c) Find $\alpha_1, \dots, \alpha_m$ (combine like terms in the factorization of $\chi_A(z)$).
     (d) Find $\gamma_1, \dots, \gamma_m$ ($\gamma_i = n - \operatorname{rank}(A - \lambda_i I)$).
  2. Find the **generalized eigenspaces** of each λ_i. This will allow us to block-diagonalize A.
     (a) For each λ_i, compute the $\ker(A - \lambda_i I) \subset \ker(A - \lambda_i I)^2 \subset \ker(A - \lambda_i I)^3 \subset \cdots$.
     (b) The sequence will stop at some $d_i \in \mathbb{N}$. In particular, it will stop when $\dim\ker(A - \lambda I)^{d_j} = \alpha_i$.

- Claim: $K_i \cap K_j = \{0\}$. Let $j_i = \dim K_i$. Take the direct sum of all K_i. Then $j_1 + \cdots + j_m = n$.

(c) Since each K_i is an invariant subspace of A, we know that there is a matrix of the linear transformation corresponding to A of the form

$$\begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_m \end{pmatrix}$$

We now just need to choose the *best* basis of each K_i, i.e., the one that makes each B_i into a (direct sum of) Jordan block(s).

3. Find the best basis for each K_i.

(a) Recall that each λ_i corresponds to $\gamma = \gamma_i$ linearly independent eigenvectors, which we will denote $v_{i,1}, \dots, v_{i,\gamma}$. We will block-diagonalize B_i into γ Jordan blocks, each of which corresponds to a $v_{i,j}$ as follows.
Every Jordan block is of the form

$$\begin{pmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix}$$

Let the above block be of dimension $k_{i,j} = d$. It follows that this block will be responsible for linearly transforming d vectors in the Jordan basis. Let $v_{i,j,1} = v_{i,j}$ be the first of these d vectors. Then the submatrix of $v_{i,j,1}$ in the Jordan basis corresponding to this Jordan block is

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

which should make sense since we want $Av_{i,j} = \lambda_i v_{i,j}$ and under this definition,

$$\begin{pmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda_i \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Now let $v_{i,j,2}$ be the second of the d vectors. Naturally, its submatrix in the Jordan basis should be

$$\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

But this implies that

$$\begin{pmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda_i \\ \vdots \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_i \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

$$Av_{i,j,2} = v_{i,j,1} + \lambda_i v_{i,j,2}$$

$$(A - \lambda I)v_{i,j,2} = v_{i,j,1}$$

Naturally, this process will generalize to show that $(A - \lambda_i I)v_{i,j,k} = v_{i,j,k-1}$, i.e., we can recursively determine the $v_{i,j,1}, \dots, v_{i,j,k_{i,j}}$.

(b) Thus, using the above process, we will find $k_{i,j}$ elements of the Jordan basis for each $v_{i,j}$. The full, ordered set of these vectors constitutes the Jordan basis.

(c) Note that each of these vectors is naturally an element of the generalized eigenspace K_i since for each $k = 1, \dots, k_{i,j}$, the formula $(A - \lambda_i I)v_{i,j,k} = v_{i,j,k-1}$ implies that

$$(A - \lambda_i I)^k v_{i,j,k} = 0$$

Also note that each $k_{i,j} \leq d_i$ and $k_{i,1} + \cdots + k_{i,\gamma} = d_i$.

- **Generalized eigenspace** (of λ): The kernel of $(A - \lambda I)^{d_\lambda}$. *Denoted by* $\boldsymbol{K_\lambda}$. *Given by*

$$K_\lambda = \ker(A - \lambda I)^{d_\lambda}$$

- $\boldsymbol{d_\lambda}$: The power of $A - \lambda I$ for which the kernel stabilizes.
- $\boldsymbol{j_\lambda}$: The dimension of the generalized eigenspace of λ. *Given by*

$$j_\lambda = \dim K_\lambda$$

- The JNF computation can be really heavy; we'll only ever compute 2×2 or 3×3 versions.
- Example:
  - Consider

$$A = \begin{pmatrix} -2 & 2 & 1 \\ -7 & 4 & 2 \\ 5 & 0 & 0 \end{pmatrix}$$

  - Then

$$\chi_A(z) = z(z-1)^2$$

  - (1) It follows that

$$\lambda_1 = 0 \qquad\qquad \lambda_2 = 1$$

  - (2) We have that

$$\ker(A - 0I) = \operatorname{span}\left\{ \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix} \right\} \qquad\qquad \ker(A - 1I) = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ -1 \\ 5 \end{pmatrix} \right\}$$

- ■ We call the left vector above q_1 and the right vector above q_2.

- Thus,
$$A \sim \left(\begin{array}{c|cc} 0 & & \\ \hline & 1 & x \\ & & 1 \end{array}\right)$$
- We find that
$$(A - 1I)^2 = \begin{pmatrix} 0 & 0 & 0 \\ 10 & -5 & -3 \\ -20 & 10 & 6 \end{pmatrix}$$
so
$$\ker(A - I)^2 = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 10 \end{pmatrix} \right\}$$
- Clearly,
$$\ker(A - I) \subsetneq \ker(A - I)^2$$
so we can stop here because the dimension of the kernel has reached the algebraic multiplicity.
- Since $q_2 \in K_1$, q_3 solves the equation $(A - I)q_3 = q_2$.
- We know that
$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} e_1 = \lambda e_1 \qquad\qquad \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} e_2 = e_1 + \lambda e_2$$
- It follows that
$$q_3 = \begin{pmatrix} 0 \\ 3 \\ -5 \end{pmatrix}$$
and hence
$$Q = \begin{pmatrix} q_1 & q_2 & q_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & -1 & 3 \\ 2 & 5 & -5 \end{pmatrix}$$
and
$$Q^{-1}AQ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

- Simple cases.
- The 2×2 case.
  - $A \in \mathcal{M}^2(\mathbb{C})$ can only have nontrivial Jordan form if it has a single eigenvalue λ with $\alpha_\lambda = 2$ and $\gamma_\lambda = 1$. If both equal 2, then $A = \lambda I_2$. If it has two eigenvalues, then it is regularly diagonalizable.
  - In this particular case, calculate λ from $\chi_Z(z) = (z - \lambda)^2$, find one eigenvector v, and find the other generalized eigenvector u; u will satisfy $(A - \lambda I)u = v$. The connecting matrix will be $Q = (v|u)$[1] and the JNF is
$$Q^{-1}AQ = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$
- The 3×3 case.
  - We divide into three nontrivial cases: $\chi_A(z) = (z - \lambda)^3$ with $\gamma_\lambda = 2$, $\chi_A(z) = (z - \lambda)^3$ with $\gamma_\lambda = 1$, and $\chi_A(z) = (z - \lambda)^2(z - \mu)$ with $\gamma_\lambda = 1$.

[1]Order matters! We need the eigenvector, specifically, to get scaled by λ only.

- In the first case, we have two eigenvectors v_1, v_2. We can find the third (generalized) eigenvector by solving $(A - \lambda I)u = v_1$ and $(A - \lambda I)u = v_2$ (only one of these will have a solution). WLOG let the first equation have a solution. Then $Q = (v_1|u|v_2)$ and the JNF is

$$Q^{-1}AQ = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}$$

- In the second case, we have one eigenvector v. We can find the second and third generalized eigenvectors by solving $(A - \lambda I)u_1 = v$ and $(A - \lambda I)u_2 = u_1$. Then $Q = (v|u_1|u_2)$ and

$$Q^{-1}AQ = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

- In the third case, we have two eigenvectors v_λ, v_μ. We can find the third (generalized) eigenvector by solving $(A - \lambda I)u = v_\lambda$. Then $Q = (v_\lambda|u|v_\mu)$ and

$$Q^{-1}AQ = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{pmatrix}$$