This document is intended as an **exact transcript** of the lecture, with extra summary and vocabulary sections for your convenience. By design, the material covered in lecture is exactly the same as what is in these notes. Due to time constraints, the notes may sometimes only contain limited illustrations, proofs, and examples; for a more thorough discussion of the course content, **consult the textbook**.

Summary

Quick summary of today's notes. Lecture starts on next page.

- Let A be an $n \times n$ matrix. Let $I = I_n$ be the $n \times n$ identity matrix.

  Let λ be a number and suppose $0 \neq v \in \mathbb{R}^n$.

  If $Av = \lambda v$ then we say that v is an *eigenvector* for A and that λ is an *eigenvalue* for A.

  More specifically, v is an *eigenvector with eigenvalue λ* for A.

- The eigenvalues of A are the solutions to the *characteristic equation* $\det(A - xI) = 0$.

  If λ is an eigenvalue then $\mathrm{Nul}(A - \lambda I)$ is the *λ-eigenspace* of A.

  To find a basis for the λ-eigenspace, use our familiar algorithm for finding bases of null spaces.

- Suppose $v_1, v_2, \dots, v_r$ are eigenvectors for A.

  Let λ_i be the eigenvalue such that $Av_i = \lambda_i v_i$.

  If $\lambda_1, \lambda_2, \dots, \lambda_r$ are all distinct, then $v_1, v_2, \dots, v_r$ are linearly independent.

- If A and B are $n \times n$ matrices and there exists an invertible $n \times n$ matrix P with

  $$A = PBP^{-1}$$

  then we say that A is *similar* to B and that B is *similar* to A.

  Any matrix is similar to itself, and if A is similar to B and B is similar to C then A is similar to C.

- Similar matrices have the same characteristic equations and same eigenvalues.

- A is *diagonalizable* if A is similar to a diagonal matrix D.

  One useful property of diagonalizable matrices: if $A = PDP^{-1}$ where D is diagonal, then there are simple formulas for each entry in the matrix $A^n = PD^nP^{-1}$ for all positive integers n.

1 Eigenvector and eigenvalues

Everywhere is this lecture, n is a positive integer and A is an $n \times n$ matrix.

Let I denote the $n \times n$ identity matrix. Let λ be a number.

Definition. A vector $v \in \mathbb{R}^n$ is an *eigenvector* for A with *eigenvalue* λ if $v \neq 0$ and $Av = \lambda v$.

The set of all $v \in \mathbb{R}^n$ with $Av = \lambda v$ is the *λ-eigenspace* of A for λ. This is just the nullspace of $A - \lambda I$.

Proposition. Let λ be a number. The following are equivalent:

1. There exists an eigenvector $v \in \mathbb{R}^n$ for A with eigenvalue λ.

   (Remember that eigenvectors must be nonzero.)
2. The matrix $A - \lambda I$ is not invertible.
3. $\det(A - \lambda I) = 0$.
4. The λ-eigenspace for A contains a nonzero vector.

As usual, a matrix is *triangular* if it is upper-triangular or lower-triangular.

The *characteristic polynomial* of a square matrix A is $\det(A - xI)$.

Theorem. The eigenvalues of a triangular square matrix A are its diagonal entries. If these numbers are $d_1, d_2, \dots, d_n$ then the characteristic polynomial of A is $(d_1 - x)(d_2 - x) \cdots (d_n - x)$.

The following is true for all square matrices, not just triangular ones.

Theorem. Suppose $\lambda_1, \lambda_2, \dots, \lambda_r$ are **distinct** eigenvalues for A, meaning $\lambda_i \neq \lambda_j$ for $i \neq j$.

Let $v_1, v_2, \dots, v_r \in \mathbb{R}^n$ be the corresponding eigenvectors, so that $Av_i = \lambda_i v_i$ for $i = 1, 2, \dots, r$.

Then the vectors $v_1, v_2, \dots v_r$ are linearly independent.

Proof. Suppose $v_1, v_2, \dots, v_r$ are linearly dependent. We argue that this leads to a logical contradiction.

There must exist an index $p > 0$ such that $v_1, v_2, \dots, v_p$ are linearly independent and v_{p+1} is a linear combination of $v_1, v_2, \dots, v_p$. (Otherwise, the vectors $v_1, v_2, \dots, v_r$ would be linearly independent.)

Let $c_1, c_2, \dots, c_p \in \mathbb{R}$ be scalars such that $v_{p+1} = c_1 v_1 + c_2 v_2 + \cdots + c_p v_p$. Then

$$\lambda_{p+1} v_{p+1} = Av_{p+1} = A(c_1 v_1 + \cdots + c_p v_p) = c_1 Av_1 + \cdots + c_p Av_p = c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_p \lambda_p v_p.$$

On the other hand, multiplying both sides of $v_{p+1} = c_1 v_1 + c_2 v_2 + \cdots + c_p v_p$ by λ_{p+1} gives

$$\lambda_{p+1} v_{p+1} = c_1 \lambda_{p+1} v_1 + c_2 \lambda_{p+1} v_2 + \cdots + c_p \lambda_{p+1} v_p.$$

By subtracting the two equations, we get

$$0 = \lambda_{p+1} v_{p+1} - \lambda_{p+1} v_{p+1} = c_1(\lambda_1 - \lambda_{p+1}) v_1 + c_2(\lambda_2 - \lambda_{p+1}) v_2 + \cdots + c_p(\lambda_p - \lambda_{p+1}) v_p.$$

Since the vectors $v_1, v_2, \dots, v_p$ are linearly independent by assumption, we must have

$$c_1(\lambda_1 - \lambda_{p+1}) = c_2(\lambda_2 - \lambda_{p+1}) = \cdots = c_p(\lambda_p - \lambda_{p+1}) = 0.$$

But the differences $\lambda_i - \lambda_{p+1}$ for $i = 1, 2, \dots, p$ are all nonzero, so we must have $c_1 = c_2 = \cdots = c_p = 0$. This implies that $v_{p+1} = 0$, contradicting our assumption that v_{p+1} is a (necessarily nonzero) eigenvector.

We conclude from this contradiction that actually the vectors $v_1, v_2, \dots, v_r$ are linearly independent. $\square$

Let x be a variable. The eigenvalues of A are precisely the solutions to the equation $\det(A - xI) = 0$ which we call the *characteristic equation* for A.

Example. The matrix

$$A = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

has characteristic polynomial $\det(A - xI) = (5-x)(3-x)(5-x)(1-x) = (5-x)^2(3-x)(1-x)$.

Since $(5-x)^2$ divides $\det(A-xI)$ but $(5-x)^3$ does not divide $\det(A-xI)$, we say that 5 is an eigenvalue of A with *algebraic multiplicity* 2. The other eigenvalues 1 and 3 have *algebraic multiplicity* 1.

In general the *algebraic multiplicity* of an eigenvalue λ for a square matrix A is the unique integer $m \geq 1$ such that $(\lambda - x)^m$ divides $\det(A - xI)$ but $(\lambda - x)^{m+1}$ does not divide $\det(A - xI)$.

We consider the following example in more depth.

Example. Consider the matrix

$$A = \begin{bmatrix} 1 & 5 & 4 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Since A is triangular, its characteristic polynomial is $(1-x)(2-x)(3-x)$ and its eigenvalues are $1, 2, 3$.

Each eigenvalue in this example has algebraic multiplicity 1. We compute the corresponding eigenspaces:

1-eigenspace. The eigenvectors of A with eigenvalue 1 are the nonzero elements of $\mathrm{Nul}(A - I)$.

$$A - I = \begin{bmatrix} 0 & 5 & 4 \\ & 1 & 0 \\ & & 2 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 \\ & 5 & 4 \\ & & 2 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 \\ & 0 & 4 \\ & & 2 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 \\ & 0 & 1 \\ & & 0 \end{bmatrix} = \mathsf{RREF}(A - I).$$

This shows that $x \in \mathrm{Nul}(A - I)$ if and only if $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, so $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is a basis for $\mathrm{Nul}(A - I)$. Therefore all eigenvectors of A with eigenvalue 1 are nonzero scalar multiples of $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

2-eigenspace. The eigenvectors of A with eigenvalue 2 are the nonzero elements of $\mathrm{Nul}(A - 2I)$.

$$A - 2I = \begin{bmatrix} -1 & 5 & 4 \\ & 0 & 0 \\ & & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -5 & 0 \\ & 0 & 1 \\ & & 0 \end{bmatrix} = \mathsf{RREF}(A - 2I).$$

This shows that $x \in \mathrm{Nul}(A - 2I)$ if and only if $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5x_2 \\ x_2 \\ 0 \end{bmatrix} = x_2 \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}$, so $\begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}$ is a basis for $\mathrm{Nul}(A - 2I)$. All eigenvectors of A with eigenvalue 2 are nonzero scalar multiples of $\begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}$.

3-eigenspace. The eigenvectors of A with eigenvalue 3 are the nonzero elements of $\mathrm{Nul}(A-3I)$.

$$A-3I=\left[\begin{array}{rrr} -2 & 5 & 4 \\ & -1 & 0 \\ 0 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{rrr} -2 & 0 & 4 \\ & 1 & 0 \\ & & 0 \end{array}\right] \sim \left[\begin{array}{rrr} 1 & 0 & -2 \\ & 1 & 0 \\ & & 0 \end{array}\right] = \mathsf{RREF}(A-3I).$$

This shows that $x \in \mathrm{Nul}(A-3I)$ if and only if $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_3 \\ 0 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ so $\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ is a basis for $\mathrm{Nul}(A-3I)$. All eigenvectors of A with eigenvalue 3 are nonzero scalar multiples of $\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$.

Since the eigenvalues 1, 2, 3, are distinct, the eigenvectors $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ are linearly independent.

Consider the **invertible** matrix whose columns are given by these linearly independent vectors:

$$P = \begin{bmatrix} 1 & 5 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

As usual, let $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. The product Pe_i is the ith column of P, so

$$Pe_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad\text{and}\quad Pe_2 = \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix} \quad\text{and}\quad Pe_3 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}.$$

Since $Px = y$ means that $P^{-1}y = P^{-1}Px = Ix = x$, it follows that

$$P^{-1}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = e_1 \quad\text{and}\quad P^{-1}\begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix} = e_2 \quad\text{and}\quad P^{-1}\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = e_3.$$

Combining these identities shows that

$$P^{-1}APe_1 = P^{-1}A\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = P^{-1}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = e_1.$$

$$P^{-1}APe_2 = P^{-1}A\begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix} = 2P^{-1}\begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix} = 2e_2.$$

$$P^{-1}APe_3 = P^{-1}A\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = 3P^{-1}\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = 3e_3.$$

These calculations determine the columns of the matrix $P^{-1}AP$.

If fact, we see that $P^{-1}AP = D$ where $D = \begin{bmatrix} e_1 & 2e_2 & 3e_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

This means that $A = P(P^{-1}AP)P^{-1} = PDP^{-1}$, so

$$\begin{bmatrix} 1 & 5 & 4 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 5 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1}.$$

One application of this decomposition: we can derive a simple formula for an arbitrary power A^n of A. Define $A^0 = I$, $A^1 = A$, $A^2 = AA$, $A^3 = AAA$, and so on.

Lemma. For any integer $n \geq 0$ we have $A^n = (PDP^{-1})^n = PD^nP^{-1}$.

Proof. Do some small examples and convince yourself that the pattern continues:

$$\begin{aligned} A^2 &= AA = PDP^{-1}PDP^{-1} = PDIDP^{-1} = PD^2P^{-1} \\ A^3 &= A^2A = PD^2P^{-1}PDP^{-1} = PD^2IDP^{-1} = PD^3P^{-1} \\ A^4 &= A^3A = PD^3P^{-1}PDP^{-1} = PD^3IDP^{-1} = PD^4P^{-1} \\ &\vdots \end{aligned}$$

and so on. □

Lemma. For any integer $n \geq 0$ we have

$$D^n = \begin{bmatrix} 1^n & 0 & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & 3^n \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & 3^n \end{bmatrix}.$$

Proof. To multiply diagonal matrices we just multiply the entries in the corresponding diagonal positions:

$$\begin{bmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_k \end{bmatrix} \begin{bmatrix} y_1 & & & \\ & y_2 & & \\ & & \ddots & \\ & & & y_k \end{bmatrix} = \begin{bmatrix} x_1y_1 & & & \\ & x_2y_2 & & \\ & & \ddots & \\ & & & x_ky_k \end{bmatrix}.$$

Therefore to evaluate $D^n = DD\cdots D$, we just raise each diagonal entry to the nth power. □

Finally, by the usual algorithm we can compute $P^{-1} = \begin{bmatrix} 1 & -5 & -2 \\ & 1 & 0 \\ & & 1 \end{bmatrix}$.

(Check that this is the correct inverse of P!)

Putting everything together gives the identity

$$\begin{aligned} A^n = PD^nP^{-1} &= \begin{bmatrix} 1 & 5 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & 3^n \end{bmatrix} \begin{bmatrix} 1 & -5 & -2 \\ & 1 & 0 \\ & & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 5 \cdot 2^n & 2 \cdot 3^n \\ 0 & 2^n & 0 \\ 0 & 0 & 3^n \end{bmatrix} \begin{bmatrix} 1 & -5 & -2 \\ & 1 & 0 \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & 5(2^n-1) & 2(3^n-1) \\ 0 & 2^n & 0 \\ 0 & 0 & 3^n \end{bmatrix}. \end{aligned}$$

Remark. We've done all these calculations for their own sake as a means of illustrating some key concepts. But these calculations would also come up in the solution of the following discrete dynamical system. Suppose $a_0, a_1, a_2, \ldots$, $b_0, b_1, b_2, \ldots$, and $c_0, c_1, c_2, \ldots$ are sequences of numbers.

For each integer $n \geq 1$, suppose

$$a_n = a_{n-1} + 5b_{n-1} + 4c_{n-1} \qquad \text{and} \qquad b_n = 2b_{n-1} \qquad \text{and} \qquad c_n = 3c_{n-1}. \tag{*}$$

How could we find a formula for a_n, b_n, and c_n in terms of n and the sequences' initial values a_0, b_0, c_0?

Note that (*) is equivalent to

$$\begin{bmatrix} a_n \\ b_n \\ c_n \end{bmatrix} = \begin{bmatrix} 1 & 5 & 4 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a_{n-1} \\ b_{n-1} \\ c_{n-1} \end{bmatrix} = A \begin{bmatrix} a_{n-1} \\ b_{n-1} \\ c_{n-1} \end{bmatrix} = A^2 \begin{bmatrix} a_{n-2} \\ b_{n-2} \\ c_{n-2} \end{bmatrix} = \dots = A^n \begin{bmatrix} a_0 \\ b_0 \\ c_0 \end{bmatrix}.$$

Thus, our formula for A^n gives

$$a_n = a_0 + 5(2^n - 1)b_0 + 2(3^n - 1)c_0 \qquad \text{and} \qquad b_n = 2^n b_0 \qquad \text{and} \qquad c_n = 3^n c_0.$$

If $a_0 = b_0 = c_0 = 1$ then $a_{10} = 123212$ and $b_{10} = 1024$ and $c_{10} = 59049$. Moreover,

$$\lim_{n\to\infty} \frac{a_n}{3^n} = \lim_{n\to\infty} \frac{a_0 + 5(2^n - 1)b_0 + 2(3^n - 1)c_0}{3^n} = 2c_0.$$

2 Similar matrices

When do square matrices have the same eigenvalues? Here is one condition that guarantees this to occur:

Definition. Two $n \times n$ matrices X and Y are *similar* if there exists an invertible $n \times n$ matrix P with

$$X = PYP^{-1}.$$

In this case it also holds that $Y = P^{-1}PYP^{-1}P = P^{-1}XP$.

If X and Y are similar, then we say that "X is *similar to* Y" and "Y is *similar to* X."

In the previous example we showed that $A = \begin{bmatrix} 1 & 5 & 4 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ and $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ are similar matrices.

There is a special name for this kind of similarity:

Definition. A square matrix X is *diagonalizable* if X is similar to a diagonal matrix

Proposition. An $n \times n$ matrix A is always similar to itself.

Proof. Since $I = I^{-1}$ we have $A = PAP^{-1}$ for $P = I$. □

Proposition. Suppose A, B, C are $n \times n$ matrices. Assume A and B are similar. Assume B and C are also similar. Then A and C are similar.

Proof. If $A = PBP^{-1}$ and $B = QCQ^{-1}$ then $R = PQ$ is invertible and $A = RCR^{-1}$. □

Theorem. If A and B are similar $n \times n$ matrices then A and B have the same characteristic polynomial and so have the same eigenvalues. (Similar matrices usually have different eigenvectors, however.)

Proof. Recall that $\det(XY) = \det(X)\det(Y)$. Assume $A = PBP^{-1}$. Then

$$A - xI = P(B - xI)P^{-1} \quad \text{and} \quad \det(A - xI) = \det(P(B - xI)P^{-1}) = \det(P)\det(B - xI)\det(P^{-1}).$$

But $\det(P)\det(P^{-1}) = \det(PP^{-1}) = \det(I) = 1$, so $\det(A - xI) = \det(B - xI)$. □

3 Vocabulary

Keywords from today's lecture:

1. **Characteristic equation** of a square matrix A.

   The equation $\det(A - xI) = 0$, where I is the identity matrix with the same size as A.

   The solutions x for this equation give all eigenvalues of A.

   Example: If $A = \begin{bmatrix} 0 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ then

   $$\det(A - xI) = \det \begin{bmatrix} -x & 2 & 0 \\ 2 & -x & 0 \\ 0 & 0 & 2-x \end{bmatrix} = (2-x)(x^2-4) = (2-x)^2(-2-x) = 0$$

   has solutions $x = 2$ and $x = -2$. These solutions are the eigenvalues for A.

2. **Algebraic multiplicity** of an eigenvalue λ of square matrix A.

   The number of times the factor $(\lambda - x)$ divides the characteristic polynomial $\det(A - xI)$.

   If $A = \begin{bmatrix} 0 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ then 2 has algebraic multiplicity 2 and -2 has algebraic multiplicity 1.

3. **Similar matrices**.

   Two $n \times n$ matrices A and B are similar if there exists an invertible $n \times n$ matrix M with

   $$A = MBM^{-1}.$$

   If A and B are similar and B and C are similar, then A and C are similar.

   Example: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ is similar to $\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}^{-1}$.

4. **Diagonalizable matrix**.

   A matrix that is similar to a diagonal matrix.