Math 115A - Week 7
Textbook sections: 4.5, 5.1-5.2
Topics covered:

- Cramer's rule
- Diagonal matrices
- Eigenvalues and eigenvectors
- Diagonalization

* * * * *

Cramer's rule

- Let A be an $n \times n$ matrix. Last week we introduced the notion of the determinant $\det(A)$ of A, and also that of a *cofactor* $\tilde{A}_{ij}$ associated to each row i and column j. Given any row i, we then have the cofactor expansion formula

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \tilde{A}_{ij}.$$

For instance, if

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

then

$$\det(A) = a\tilde{A}_{11} - b\tilde{A}_{12} + c\tilde{A}_{13}$$

or in other words

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \det \begin{pmatrix} d & e \\ g & h \end{pmatrix}.$$

- Suppose we replace the row (a, b, c) by (d, e, f) in the above example. Then we have

$$\det\begin{pmatrix} d & e & f \\ d & e & f \\ g & h & i \end{pmatrix} = d\det\begin{pmatrix} e & f \\ h & i \end{pmatrix} - e\det\begin{pmatrix} d & f \\ g & i \end{pmatrix} + f\det\begin{pmatrix} d & e \\ g & h \end{pmatrix}.$$

  But the left-hand side is zero because two of the rows are the same (see Property 1 of determinants on the previous week's notes). Thus we have

$$0 = d\tilde{A}_{11} - e\tilde{A}_{12} + f\tilde{A}_{13}.$$

  Similarly we have

$$0 = g\tilde{A}_{11} - h\tilde{A}_{12} + i\tilde{A}_{13}.$$

  We can also do the same analysis with the cofactor expansion along the second row

$$\det(A) = -d\tilde{A}_{21} + e\tilde{A}_{22} - f\tilde{A}_{23}$$

  yielding

$$0 = -a\tilde{A}_{21} + b\tilde{A}_{22} - c\tilde{A}_{23}$$

$$0 = -g\tilde{A}_{21} + h\tilde{A}_{22} - i\tilde{A}_{23}.$$

  And similarly for the third row:

$$\det(A) = g\tilde{A}_{31} - h\tilde{A}_{32} + i\tilde{A}_{33}$$

$$0 = a\tilde{A}_{31} - b\tilde{A}_{33} + c\tilde{A}_{33}$$

$$0 = d\tilde{A}_{31} - e\tilde{A}_{32} + f\tilde{A}_{33}.$$

  We can put all these nine identies together in a compact matrix form as

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}\begin{pmatrix} +\tilde{A}_{11} & -\tilde{A}_{21} & +\tilde{A}_{31} \\ -\tilde{A}_{12} & +\tilde{A}_{22} & -\tilde{A}_{32} \\ +\tilde{A}_{13} & -\tilde{A}_{23} & +\tilde{A}_{33} \end{pmatrix} = \begin{pmatrix} \det(A) & 0 & 0 \\ 0 & \det(A) & 0 \\ 0 & 0 & \det(A) \end{pmatrix}.$$

The second matrix on the left-hand side is known as the *adjugate* of A, and is denoted $adj(A)$:

$$adj(A) := \begin{pmatrix} \tilde{A}_{11} & -\tilde{A}_{21} & \tilde{A}_{31} \\ -\tilde{A}_{12} & \tilde{A}_{22} & -\tilde{A}_{32} \\ \tilde{A}_{13} & -\tilde{A}_{23} & \tilde{A}_{33} \end{pmatrix}.$$

Thus we have the identity for 3×3 matrices

$$A adj(A) = \det(A) I_3.$$

The adjugate matrix is the transpose of the *cofactor matrix* $cof(A)$:

$$cof(A) := \begin{pmatrix} \tilde{A}_{11} & -\tilde{A}_{12} & \tilde{A}_{13} \\ -\tilde{A}_{21} & \tilde{A}_{22} & -\tilde{A}_{23} \\ \tilde{A}_{31} & -\tilde{A}_{32} & \tilde{A}_{33} \end{pmatrix}.$$

Thus $adj(A) = cof(A)^t$. To compute the cofactor matrix, at every row i and column j we extract the minor corresponding to that row and column, take the $n-1 \times n-1$ determinant of the minor, and then place that number in the ij entry of the cofactor matrix. Then we alternate the signs by $(-1)^{i+j}$.

- More generally, for $n \times n$ matrices, we can define the cofactor matrix by

$$cof(A)_{ij} = (-1)^{i+j} \tilde{A}_{ij}$$

and the adjugate matrix by $adj(A) = cof(A)^t$, so that

$$adj(A)_{ij} = (-1)^{j+i} \tilde{A}_{ji},$$

and then we have the identity

$$A adj(A) = \det(A) I_n.$$

If $\det(A)$ is non-zero, then A is invertible and we thus have

$$A^{-1} = \frac{1}{\det(A)} adj(A).$$

This is known as *Cramer's rule* - it allows us to compute the inverse of a matrix using determinants. (There is also a closely related rule, also known as Cramer's rule, which allows one to solve equations $Ax = b$ when A is invertible; see the textbook).

- For example, in the 2×2 case

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

then the cofactor matrix is

$$cof(A) = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

and so the adjugate matrix is

$$adj(A) = cof(A)^t = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

and so, if $\det(A) \neq 0$, the inverse of A is

$$A^{-1} = \frac{1}{\det(A)} adj(A) = \frac{1}{ad - bc} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}.$$

* * * * *

Diagonal matrices

- Matrices in general are complicated objects to manipulate, and we are always looking for ways to simplify them into something better. Last week we explored one such way to do so: using elementary row (or column operations) to reduce a matrix into row-echelon form, or to even simpler forms. This type of simplification is good for certain purposes (computing rank, determinant, inverse), but is not good for other purposes. For instance, suppose you want to raise a matrix A to a large power, say A^{100}. Using elementary matrices to reduce A to, say, row-echelon form will not be very helpful, because (a) it is still not very easy to raise row-echelon form matrices to very large powers, and (b) one has to somehow deal with all the elementary matrices you used to convert A into row echelon form. However, to perform tasks like this there is a better factorization available, known as *diagonalization*. But before we do this, we first digress on diagonal matrices.

- **Definition** An $n \times n$ matrix A is said to be *diagonal* if all the off-diagonal entries are zero, i.e. $A_{ij} = 0$ whenever $i \neq j$. Equivalently, a diagonal matrix is of the form

$$A = \begin{pmatrix} A_{11} & 0 & \ldots & 0 \\ 0 & A_{22} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & A_{nn} \end{pmatrix}.$$

We write this matrix as $\mathrm{diag}(A_{11}, A_{22}, \ldots, A_{nn})$. Thus for instance

$$\mathrm{diag}(1, 3, 5) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

- Diagonal matrices are very easy to add, scalar multiply, and multiply. One can easily verify that

$$\mathrm{diag}(a_1, a_2, \ldots, a_n) + \mathrm{diag}(b_1, b_2, \ldots, b_n) = \mathrm{diag}(a_1+b_1, a_2+b_2, \ldots, a_n+b_n),$$

$$c\,\mathrm{diag}(a_1, a_2, \ldots, a_n) = \mathrm{diag}(ca_1, ca_2, \ldots, ca_n)$$

and

$$\mathrm{diag}(a_1, a_2, \ldots, a_n)\mathrm{diag}(b_1, b_2, \ldots, b_n) = \mathrm{diag}(a_1 b_1, a_2 b_2, \ldots, a_n b_n).$$

Thus for instance

$$\mathrm{diag}(1, 3, 5)\mathrm{diag}(1, 3, 5) = \mathrm{diag}(1^2, 3^2, 5^2)$$

and more generally

$$\mathrm{diag}(1, 3, 5)^n = \mathrm{diag}(1^n, 3^n, 5^n).$$

Thus raising diagonal matrices to high powers is very easy. More generally, polynomial expressions of a diagonal matrix are very easy to compute. For instance, consider the polynomial $f(x) = x^3 + 4x^2 + 2$. We can apply this polynomial to any $n \times n$ matrix A, creating the new matrix $f(A) = A^3 + 4A^2 + 2$. In general, such a matrix may be difficult to compute. But for a diagonal matrix $A = \mathrm{diag}(a_1, a_2, \ldots, a_n)$, we have

$$A^2 = \mathrm{diag}(a_1^2, a_2^2, \ldots, a_n^2)$$
$$A^3 = \mathrm{diag}(a_1^3, a_2^3, \ldots, a_n^3)$$

and thus

$$f(A) = \mathrm{diag}(a_1^3+4a_1^2+2, a_2^3+4a_2^2+2, \ldots, a_n^3+4a_n^2+2) = \mathrm{diag}(f(a_1), \ldots, f(a_n)).$$

This is true for more general polynomials f:

$$f(\mathrm{diag}(a_1, \ldots, a_n)) = \mathrm{diag}(f(a_1), \ldots, f(a_n)).$$

Thus to do any sort of polynomial operation to a diagonal matrix, one just has to perform it on the diagonal entries separately.

- If A is an $n \times n$ diagonal matrix $A = \mathrm{diag}(a_1, \ldots, a_n)$, then the linear transformation $L_A : \mathbf{R}^n \to \mathbf{R}^n$ is very simple:

$$L_A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_1 & 0 & \ldots & 0 \\ 0 & a_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_1 x_1 \\ a_2 x_2 \\ \vdots \\ a_n x_n \end{pmatrix}.$$

  Thus L_A dilates the first co-ordinate x_1 by a_1, the second co-ordinate x_2 by a_2, and so forth. In particular, if $\beta = (e_1, e_2, \ldots, e_n)$ is the standard ordered basis for $\mathbf{R}^n$, then

$$L_A e_1 = a_1 e_1; \quad L_A e_2 = a_2 e_2; \ldots; \quad L_A e_n = a_n e_n.$$

  This leads naturally to the concept of *eigenvalues* and *eigenvalues*, which we now discuss.

- Remember from last week that the rank of a matrix is equal to the number of non-zero rows in row echelon form. Thus it is easy to see that

- **Lemma 1.** The rank of a diagonal matrix is equal to the number of its non-zero entries.

- Thus, for instance, $\mathrm{diag}(3,4,5,0,0,0)$ has rank 3.

* * * * *

Eigenvalues and eigenvectors

- Let $T : V \to V$ be a linear transformation from a vector space V to itself. One of the simplest possible examples of such a transformation is the identity transformation $T = I_V$, so that $Tv = v$ for all $v \in V$. After the identity operation, the next simplest example of such a transformation is a dilation $T = \lambda I_V$ for some scalar λ, so that $Tv = \lambda v$ for all $v \in V$.

- In general, though, T does not look like a dilation. However, there are often some special vectors in V for which T is as simple as a dilation, and these are known as *eigenvectors*.

- **Definition** An *eigenvector* v of T is a non-zero vector $v \in V$ such that $Tv = \lambda v$ for some scalar λ. The scalar λ is known as the *eigenvalue* corresponding to v.

- **Example** Consider the linear transformation $T : \mathbf{R}^2 \to \mathbf{R}^2$ defined by $T(x,y) := (5x, 3y)$. Then the vector $v = (1,0)$ is an eigenvector of T with eigenvalue 5, since $Tv = T(1,0) = (5,0) = 5v$. More generally, any non-zero vector of the form $(x,0)$ is an eigenvector with eigenvalue 5. Similarly, $(0,y)$ is an eigenvector of T with eigenvalue 3, if y is non-zero. The vector $v = (1,1)$ is *not* an eigenvector, because $Tv = (5,3)$ is not a scalar multiple of v.

- **Example** More generally, if $A = \mathrm{diag}(a_1, \ldots, a_n)$ is a diagonal matrix, then the basis vectors $e_1, \ldots, e_n$ are eigenvectors for L_A, with eigenvalues $a_1, \ldots, a_n$ respectively.

- **Example** If T is the identity operator, then every non-zero vector is an eigenvector, with eigenvalue 1 (why?). More generally, if $T = \lambda I_V$ is λ times the identity operator, then every non-zero vector is an eigenvector, with eigenvalue λ (Why?).

- **Example** If $T : V \to V$ is any linear transformation, and v is any non-zero vector in the null space $N(T)$, then v is an eigenvector with eigenvalue 0. (Why?)

- **Example** Let $T : \mathbf{R}^2 \to \mathbf{R}^2$ be the reflection through the line l connecting the origin to $(4, 3)$. Then $(4, 3)$ is an eigenvector with eigenvalue 1 (why?), and $(3, -4)$ is an eigenvector with eigenvalue -1 (why?).

- We do not consider the 0 vector as an eigenvector, even though $T0$ is always 0, because we cannot determine what eigenvalue 0 should have.

- If A is an $n \times n$ matrix, we say that v is an *eigenvector* for A with eigenvalue λ if it is already an eigenvector for L_A with eigenvalue λ, i.e. $Av = \lambda v$. In other words, for the purposes of computing eigenvalues and eigenvectors we do not distinguish between a matrix A and its linear transformation L_A.

- (Incidentally, the word "eigen" is German for "own". An eigenvector is a vector which keeps its *own* direction when acted on by T. The terminology is thus a hybrid of German and English, though some people prefer "principal value" and "principal vector" to avoid this (or "characteristic" or "proper" instead of "principal"). Then again, "vector" is pure Latin. English is very cosmopolitan).

- **Definition** Let $T : V \to V$ be a linear transformation, and let λ be a scalar. Then the *eigenspace* of T corresponding to λ is the set of all vectors (including 0) such that $Tv = \lambda v$.

- Thus an eigenvector with eigenvalue λ is the same thing as a non-zero element of the eigenspace with eigenvalue λ. Since $Tv = \lambda v$ is equivalent to $(T - \lambda I_V)v = 0$, we thus see that the eigenspace of T with eigenvalue λ is the same thign as the null space $N(T - \lambda I_V)$ of $T - \lambda I_V$. In particular, the eigenspace is always a subspace of V. From the above discussion we also see that λ is an eigenvalue of T if and only if $N(T - \lambda I_V)$ is non-zero, i.e. when $T - \lambda I_V$ is not one-to-one.

- **Example** Let $T : \mathbf{R}^2 \to \mathbf{R}^2$ be the transformation $T(x, y) := (5x, 3y)$. Then the x-axis is the eigenspace $N(T - 3I_{\mathbf{R}^2})$ with eigenvalue 5, while the y-axis is the eigenspace $N(T - 5I_{\mathbf{R}^2})$ with eigenvalue 5. For all other values $\lambda \neq 3, 5$, the eigenspace $N(T - \lambda I_{\mathbf{R}^2})$ is just the zero vector space $\{0\}$ (why?).

- The relationship between eigenvectors and diagonal matrices is the following.

- **Lemma 2.** Suppose that V is an n-dimensional vector space, and suppose that $T : V \to V$ is a linear transformation. Suppose that V has an ordered basis $\beta = (v_1, v_2, \ldots, v_n)$, such that each v_j is an eigenvector of T with eigenvalue λ_j. Then the matrix $[T]_\beta^\beta$ is a diagonal matrix; in fact $[T]_\beta^\beta = \text{diag}(\lambda_1, \ldots, \lambda_n)$.

- Conversely, if $\beta = (v_1, v_2, \ldots, v_n)$ is a basis such that $[T]_\beta^\beta = \text{diag}(\lambda_1, \ldots, \lambda_n)$, then each v_j is an eigenvector of T with eigenvalue λ.

- **Proof** Suppose that v_j is an eigenvector of T with eigenvalue λ_j. Then $Tv_j = \lambda_j v_j$, so $[Tv_j]^\beta$ is just the column vector with j^{th} entry equal to λ_j, and all other entries zero. Putting all these column vectors together we see that $[T]_\beta^\beta = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Conversely, if $[T]_\beta^\beta = \text{diag}(\lambda_1, \ldots, \lambda_n)$, then by definition of $[T]_\beta^\beta$ we see that $Tv_j = \lambda_j v_j$, and so v_j is an eigenvector with eigenvalue λ_j. □

- **Definition.** A linear transformation $T : V \to V$ is said to be *diagonalizable* if there is an ordered basis β of V for which the matrix $[T]_\beta^\beta$ is diagonal.

- Lemma 2 thus says that a transformation is diagonalizable if and only if it has a basis consisting entirely of eigenvectors.

- **Example** Let $T : \mathbf{R}^2 \to \mathbf{R}^2$ be the reflection through the line l connecting the origin to $(4, 3)$. Then $(4, 3)$ and $(3, -4)$ are both eigenvectors for T. Since these two vectors are linearly independent and $\mathbf{R}^2$ is two-dimensional, they form a basis for T. Thus T is diagonalizable; indeed, if $\beta := ((4, 3), (3, -4))$, then

$$[T]_\beta^\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \text{diag}(1, -1).$$

- If one knows how to diagonalize a transformation, then it becomes very easy to manipulate. For example, in the above reflection example we see very quickly that T must have rank 2 (since $\mathrm{diag}(1,-1)$ has two non-zero entries). Also, we can square T easily:

$$[T^2]_\beta^\beta = \mathrm{diag}(1,-1)^2 = \mathrm{diag}(1,1) = I_2 = [I_{\mathbf{R}^2}]_\beta^\beta$$

and hence $T^2 = I_{\mathbf{R}^2}$, the identity transformation. (Geometrically, this amounts to the fact that if you reflect twice around the same line, you get the identity).

- **Definition.** An $n \times n$ matrix A is said to be *diagonalizable* if the corresponding linear transformation L_A is diagonalizable.

- **Example** The matrix $A = \mathrm{diag}(5,3)$ is diagonalizable, because the linear operator L_A in the standard basis $\beta = ((1,0),(0,1))$ is just A itself: $[L_A]_\beta^\beta = A$, which is diagonal. So all diagonal matrices are diagonalizable (no surprise there).

- **Lemma 3.** A matrix A is diagonalizable if and only if $A = QDQ^{-1}$ for some invertible matrix Q and some diagonal matrix D. In other words, a matrix is diagonalizable if and only if it is similar to a diagonal matrix.

- **Proof** Suppose A was diagonalizable. Then $[L_A]_{\beta'}^{\beta'}$ would be equal to some diagonal matrix D, for some choice of basis β' (which may be different from the standard ordered basis β. But by the change of variables formula,

  $$A = [L_A]_\beta^\beta = Q[L_A]_{\beta'}^{\beta'}Q^{-1} = QDQ^{-1}$$

  as desired, where $Q := [I_{\mathbf{R}^n}]_{\beta'}^{\beta}$ is the change of variables matrix.

  Conversely, suppose that $A = QDQ^{-1}$ for some invertible matrix Q. Write $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, so that $De_j = \lambda_j$. Then

  $$A(Qe_j) = QDQ^{-1}Qe_j = QDe_j = Q\lambda_j e_j = \lambda_j(Qe_j)$$

  and so Qe_j is an eigenvector for A with eigenvalue λ_j. Since Q is invertible and $e_1, \ldots, e_n$ is a basis, we see that $Qe_1, \ldots, Qe_n$ is also a basis (why?). Thus we have found a basis of $\mathbf{R}^n$ consisting entirely of eigenvectors of A, and so A is diagonalizable. $\square$

- From Lemma 2 and Lemma 3 we see that if we can find a basis $(v_1, \ldots, v_n)$ of $\mathbf{R}^n$ which consists entirely of eigenvectors of A, then A is diagonalizable and $A = QDQ^{-1}$ for some diagonal matrix D and some invertible matrix Q. We now make this statement more precise, specifying precisely what Q and D are.

- **Lemma 4.** Let A be an $n \times n$ matrix, and suppose that $(v_1, \ldots, v_n)$ is an ordered basis of $\mathbf{R}^n$ such that each v_j is an eigenvector of A with eigenvalue λ_j (i.e. $Av_j = \lambda v_j$ for $j = 1, \ldots, n$). Then we have

$$A = Q\mathrm{diag}(\lambda_1, \ldots, \lambda_n)Q^{-1}$$

where Q is the $n \times n$ matrix with columns $v_1, v_2, \ldots, v_n$:

$$Q = (v_1, v_2, \ldots, v_n).$$

- **Proof.** Let β' be the ordered basis $\beta' := (v_1, v_2, \ldots, v_n)$ of $\mathbf{R}^n$, and let $\beta := (e_1, e_2, \ldots, e_n)$ be the standard ordered basis of $\mathbf{R}^n$. Then

$$[I_{\mathbf{R}^n}]_{\beta'}^{\beta} = Q$$

(why?). So by the change of variables formula

$$A = [L_A]_{\beta}^{\beta} = Q[L_A]_{\beta'}^{\beta'}Q^{-1}.$$

On the other hand, since $L_A v_j = \lambda_j v_j$, we see that

$$[L_A]_{\beta'}^{\beta'} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

Combining these two equations we obtain the lemma. □.

* * * * *

Computing eigenvalues

- Now we compute the eigenvalues and eigenvectors of a general matrix. The key lemma here is

- **Lemma 5.** A scalar λ is an eigenvalue of an $n \times n$ square matrix A if and only $\det(A - \lambda I_n) = 0$.

- **Proof.** If λ is an eigenvalue of A, then $Av = \lambda v$ for some non-zero v, thus $(A - \lambda I_n)v = 0$. Thus $A - \lambda I_n$ is not invertible, and so $\det(A - \lambda I_n) = 0$ (Theorem 12 from last week's notes). Conversely, if $\det(A - \lambda I_n) = 0$, then $A - \lambda I_n$ is not invertible (again by Theorem 12), which means that the corresponding linear transformation is not one-to-one (recall that one-to-one and onto are equivalent when the domain and range have the same dimension; see Lemma 2 of Week 3 notes). So we have $(A - \lambda I_n)v = 0$ for some non-zero v, which means that $Av = \lambda v$ and hence λ is an eigenvalue. □

- Because of this lemma, we call $\det(A - \lambda I_n)$ the *characteristic polynomial* of A, and sometimes call it $f(\lambda)$. Lemma 5 then says that the eigenvalues of A are precisely the zeroes of $f(\lambda)$.

- **Example** Let A be the matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

  Then the characteristic polynomial $f(\lambda)$ is given by

$$f(\lambda) = \det \begin{pmatrix} -\lambda & 1 \\ 1 & 1-\lambda \end{pmatrix} = -\lambda(1-\lambda) - 1 \times 1 = \lambda^2 - \lambda - 1.$$

  From the quadratic formula, this polynomial has zeroes when $\lambda = (1 \pm \sqrt{5})/2$, and so the eigenvalues are $\lambda_1 := (1+\sqrt{5})/2 = 1.618\ldots$ and $\lambda_2 = (1-\sqrt{5})/2 := -0.618\ldots$.

- Once we have the eigenvalues of A, we can compute eigenvectors, because the eigenvectors with eigenvalues λ are precisely those non-zero vectors in the null-space of $A - \lambda I_n$ (or equivalently, of the null space of $L_A - \lambda I_{\mathbf{R}^n}$).

- **Example:** Let A be the above matrix. Let us try to find the eigenvectors $\begin{pmatrix} x \\ y \end{pmatrix}$ with eigenvalue $\lambda_1 = (1+\sqrt{5})/2$. In other words, we want to solve the equation

$$(A - (1+\sqrt{5})/2) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

or in other words

$$\begin{pmatrix} -(1+\sqrt{5})/2 & 1 \\ 1 & (1-\sqrt{5})/2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

or equivalently

$$-\frac{1+\sqrt{5}}{2}x + y = 0$$

$$x + \frac{1-\sqrt{5}}{2}y = 0.$$

This matrix does not have full rank (since its determinant is zero - why should this not be surprising?). Indeed, the second equation here is just $(1-\sqrt{5})/2$ times the first equation. So the general solution is y arbitrary, and x equal to $-\frac{1-\sqrt{5}}{2}y$. In particular, we have

$$v_1 := \begin{pmatrix} -\frac{1-\sqrt{5}}{2} \\ 1 \end{pmatrix}$$

as an eigenvector of A with eigenvalue $\lambda_1 = (1+\sqrt{5})/2$.

A similar argument gives

$$v_2 := \begin{pmatrix} -\frac{1+\sqrt{5}}{2} \\ 1 \end{pmatrix}$$

as an eigenvector of A with eigenvalue $\lambda_2 = (1-\sqrt{5})/2$. Thus we have $Av_1 = \lambda_1 v_1$ and $Av_2 = \lambda_2 v_2$. Thus, if we let β' be the ordered basis $\beta' := (v_1, v_2)$, then

$$[L_A]_{\beta'}^{\beta'} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Thus A is diagonalizable. Indeed, from Lemma 4 we have

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

where $D := \text{diag}(\lambda_1, \lambda_2)$ and $Q := (v_1, v_2)$.

- As an application, we recall the example of Fibonacci's rabbits from Week 2. If at the beginning of a year there are x pairs of juvenile rabbits and y pairs of adult rabbits - which we represent by the vector $\begin{pmatrix} x \\ y \end{pmatrix}$ in $\mathbf{R}^2$ - then at the end of the next year there will be y juvenile pairs and $x+y$ adult pairs - so the new vector is

$$\begin{pmatrix} y \\ x+y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix}.$$

Thus, each passage of a year multiplies the population vector by A. So if we start with one juvenile pair and no adult pairs - so the population vector is initially $v_0 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ - then after n years, the population vector should become $A^n v_0$. To compute this, one would have to multiply A by itself n times, which appears to be difficult (try computing A^5 by hand, for instance!). However, this can be done efficiently using the diagonalization $A = QDQ^{-1}$ we have. Observe that

$$A^2 = QDQ^{-1}QDQ^{-1} = QD^2Q^{-1}$$

$$A^3 = A^2A = QD^2Q^{-1}QDQ^{-1} = QD^3Q^{-1}$$

and more generally (by induction on n)

$$A^n = QD^nQ^{-1}.$$

In particular, our population vector after n years is

$$A^n v_0 = QD^nQ^{-1}v_0.$$

But since D is the diagonal matrix $D = \text{diag}(\lambda_1, \lambda_2)$, D^n is easy to compute:

$$D^n = \text{diag}(\lambda_1^n, \lambda_2^n).$$

Now we can compute $QD^nQ^{-1}v_0$. Since

$$Q = (v_1, v_2) = \begin{pmatrix} -\frac{1-\sqrt{5}}{2} & -\frac{1+\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}$$

we have

$$\det(Q) = -\frac{1-\sqrt{5}}{2} + \frac{1-\sqrt{5}}{2} = \sqrt{5}$$

and so by Cramer's rule

$$Q^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & \frac{1+\sqrt{5}}{2} \\ -1 & -\frac{1-\sqrt{5}}{2} \end{pmatrix}$$

and so

$$Q^{-1} v_0 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

and hence

$$D^n Q^{-1} v_0 = \operatorname{diag}(\lambda_1^n, \lambda_2^n) \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda_1^n \\ -\lambda_2^n \end{pmatrix}.$$

Since $\frac{1-\sqrt{5}}{2} \frac{1+\sqrt{5}}{2} = \frac{1-5}{4} = -1$, we have

$$Q = \begin{pmatrix} -\frac{1-\sqrt{5}}{2} & -\frac{1+\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \lambda_1^{-1} & \lambda_2^{-1} \\ 1 & 1 \end{pmatrix}$$

and hence

$$A^n v_0 = Q D^n Q^{-1} v_0 = \begin{pmatrix} \lambda_1^{-1} & \lambda_2^{-1} \\ 1 & 1 \end{pmatrix} \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda_1^n \\ -\lambda_2^n \end{pmatrix} = \begin{pmatrix} (\lambda_1^{n-1} - \lambda_2^{n-1})/\sqrt{5} \\ (\lambda_1^n - \lambda_2^n)/\sqrt{5} \end{pmatrix}.$$

Thus, after n years, the number of pairs of juvenile rabbits is

$$F_{n-1} = (\lambda_1^{n-1} - \lambda_2^{n-1})/\sqrt{5} = ((1.618\ldots)^{n-1} - (-0.618\ldots)^{n-1})/2.236\ldots,$$

and the number of pairs of adult rabbits is

$$F_n = (\lambda_1^n - \lambda_2^n)/\sqrt{5} = ((1.618\ldots)^n - (-0.618\ldots)^n)/2.236\ldots.$$

This is a remarkable formula - it does not look like it at all, but the expressions F_{n-1}, F_n are always integers. For instance

$$F_3 = ((1.618\ldots)^3 - (-0.618\ldots)^3)/2.236\ldots = 2.$$

(Check this!). The numbers

$$F_0 = 0, F_1 = 1, F_2 = 1, F_3 = 2, F_4 = 3, F_5 = 5, F_6 = 8, F_7 = 13, F_8 = 25, \ldots$$

are known as *Fibonacci numbers* and come up in all sorts of places (including, oddly enough, the number of petals on flowers and pine cones). The above formula shows that these numbers grow exponentially, and are comparable to $(1.618)^n$ when n gets large. The number $1.618\ldots = (1+\sqrt{5})/2$ is known as the *golden ratio* and has several interesting properties, which we will not go into here.

- **A final note.** Up until now, we have always chosen the field of scalars to be real. However, it will now sometimes be convenient to change the field of scalars to be complex, because one gets more eigenvalues and eigenvectors this way. For instance, consider the matrix

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

The characteristic polynomial $f(\lambda)$ is given by

$$f(\lambda) = \det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1.$$

If one restricts the field of scalars to be real, then $f(\lambda)$ has no zeroes, and so there are no real eigenvalues (and thus no real eigenvectors). On the other hand, if one expands the field of scalars to be complex, then $f(\lambda)$ has zeroes at $\lambda = \pm i$, and one can easily show that vectors such as $\begin{pmatrix} 1 \\ i \end{pmatrix}$ are eigenvectors with eigenvalue i, while $\begin{pmatrix} 1 \\ i \end{pmatrix}$ is an eigenvector with eigenvalue $-i$. Thus it is sometimes advantageous to introduce complex numbers into a problem which seems purely concerned with real numbers, because it can introduce such useful concepts as eigenvectors and eigenvalues into the situation. (An example of this appears in Q10 of this week's assignment).