NOTES FOR 10 OCT (THURSDAY)

1. Recap

(1) Proved all the familiar formulae about determinants, including Cramer's rule and a formula for the inverse.

2. Eigenvalues and Eigenvectors

Consider these two problems :

(1) Given a matrix R that represents a rotation in $\mathbb{R}^3$, find its axis of rotation.
(2) Solve $\frac{dx}{dt} = 2x + 3y$, $\frac{dy}{dt} = 3x + 2y$.
(3) The chance of it raining tomorrow if it rains today is 0.7 and if it does not rain, it is 0.5. Given that it rained today, what is the chance of it raining after many days ?

We already discussed the solutions of the first and the third.

(1) Change the variables $z_1 = x + y$, $z_2 = x - y$. Then $\frac{dz_1}{dt} = 5z_1$, $\frac{dz_2}{dt} = -z_2$ whose solution is $z_1 = ae^{5t}, z_2 = be^{-t}$. Hence, $x = \frac{ae^{5t}-be^{-t}}{2}$, $y = \frac{ae^{5t}-be^{-t}}{2}$. (Morally, $d\vec{v}dt = A\vec{v}$. So we expect $\vec{v} = e^{At}\vec{v}_0$, whatever e^{At} means.)

The above problems suggest that given an $n \times n$ matrix A, finding vectors v such that $Av = \lambda v$ is helpful. It is even more helpful if one can "change variables" to v, i.e., if the v's so obtained form a basis. So we define : Given a linear map $T : V \to V$ where V is a finite-dimensional vector space over a field $\mathbb{F}$, a non-zero vector $v \in V$ is called an eigenvector of T with eigenvalue λ if $Tv = \lambda v$.

Lemma 2.1. *λ is an eigenvalue of T iff it is a root of the characteristic polynomial* $\det(\lambda I - T)$.

Proof. λ is an eigenvalue of T iff there exists a non-zero $v \in V$ such that $Tv = \lambda v$ iff $(T - \lambda I)v = 0$ iff $T - \lambda I$ is singular iff $\det(\lambda I - T) = 0$. □

It is easy to see that the collection of vectors $v \in V$ such that $Tv = \lambda v$ is a subspace of V (whose dimension is > 0 iff λ is an eigenvalue). When λ is an eigenvalue, the subspace is called the "eigenspace of λ". The dimension of the eigenspace of λ is called the geometric multiplicity of λ. More generally, a subspace $W \subset V$ is said to be an invariant subspace if $T(W) \subset W$. Here is an observation : If $e_1, \ldots, e_w$ is a basis for W and $e_{w+1}, \ldots, e_n$ extends it to V, then the matrix of T in this basis is $\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$. Here are examples/non-examples.

(1) Let $T : \mathbb{Z}_3^2 \to \mathbb{Z}_3^2$ be the linear map $Tx = Ax$ where $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$. The characteristic polynomial is $(\lambda - 1)(\lambda - 2) - 1$. It is never zero. So no eigenvalues.
(2) $T : \mathbb{Q}^2 \to \mathbb{Q}^2$ such that $Tx = Ax$ where A is as above. Even here there are no eigenvalues. If we allow reals, then $\lambda = \frac{3 \pm \sqrt{5}}{2}$.

(3) However, if $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then there are no real eigenvalues. There are complex ones though.
(4) Let $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ where $a, b, c \in \mathbb{R}$. Then $p(\lambda) = \lambda^2 - (a+c)\lambda + ac - b^2$. Firstly, note that $a + c = tr(A)$ and $ac - b^2 = \det(A)$. Secondly, this matrix has two (not necessarily distinct) real eigenvalues regardless of a, b, c.
(5) Let V be the space of all abstract complex polynomials. Then $T : V \to V$ given by $T(p) = xp$ has no eigenvectors or eigenvalues.

These examples/non-examples suggest that

(1) Eigenvalues need not exist in $\mathbb{F}$. However, by the fundamental theorem of algebra, an $n \times n$ matrix with complex entries has n eigenvalues (when counted with multiplicity). In particular, there is at least one eigenvector. Here is a definition : If $p(\lambda) = (\lambda - \lambda_1)^{a_1}(\lambda - \lambda_2)^{a_2} \ldots$, then λ_i is said to have algebraic multiplicity equal to a_i. Clearly, $a_i \leq n$ and $\sum_i a_i = n$.
(2) Infinite-dimensions are badly behaved with regard to eigenvalues and vectors.
(3) The coefficients of the characteristic polynomial can potentially tell us about the trace and determinant. Actually,

Lemma 2.2. *Let $\mathbb{F}$ be any field and let $A \in Mat_{n\times n}(\mathbb{F})$. The characteristic polynomial is invariant under similarity. Moreover, the constant term is $(-1)^n \det(A)$ and the coefficient of λ^{n-1} is $-tr(A)$. Also, A and A^T have the same characteristic polynomials.*

Proof. $\det(\lambda I - PAP^{-1}) = \det(P(\lambda I - A)P^{-1}) = p_A(\lambda)$. Also, $p_{A^T}(\lambda) = \det((\lambda I - A)^T) = p_A(\lambda)$. $p(0) = \det(-A) = (-1)^n \det(A)$. We prove the trace equality inductively (on n). The $n = 1$ case is trivial. Assuming truth for $1, 2 \ldots, n-1$, $\det(\lambda I - A) = (\lambda - a_{11})M_{11} - a_{12}M_{12} \ldots$. So $coeff(\lambda^{n-1})$ in $p_A(\lambda)$ is $-a_{11}coeff(\lambda^{n-1})$ in M_{11} plus similar terms for a_{12} etc plus $-tr([A]_{(n-1)\times(n-1)})$ inductively. We are done because terms like M_{12} do not have λ^{n-1}. □

(4) Real symmetric matrices might have real eigenvalues. We will return to this observation much later.

Going back to the motivating problems above, we make a definition : A linear map $T : V \to V$ is called diagonalisable if there is a basis of V consisting of eigenvectors. Note that in such a basis, if V is finite-dimensional, the matrix of T is diagonal (D) with diagonal entries being the eigenvalues. If the matrix is A in some other basis and $A\vec{v}_{some\ other} = \vec{v}_{eigenvector\ basis}$, then $D = PAP^{-1}$ where the i^{th} column of P^{-1} is the i^{th} eigenvector (in some order) written in the "some other" basis. Taking cue from this, a matrix A is called diagonalisable if it is similar to a diagonal matrix. (Note that diagonal matrices are of course diagonalisable.)

The point is if $A = P^{-1}DP$, then $A^k = P^{-1}D^kP$. So powers can be calculated easily. Here are examples/non-examples.

(1) The matrix $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ has only one eigenvalue : 0. It's algebraic multiplicity is 2 and geometric multiplicity is 1 (spanned by $(1, 0)$). It is clearly not diagonalisable (only one linearly independent eigenvector). Nonetheless, it is in upper-triangular form. Also, $p(\lambda) = \lambda^2$ and $p(A) = A^2 = [0]_{2\times 2}$. Curious !
(2) The matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ has two (not neccesarily distinct) real eigenvalues if $a, b, c \in \mathbb{R}$. An explicit calculation shows that it is diagonalisable. Moreover, $p(A) = (A - \lambda_1)(A - \lambda_2) =$

$P(D-\lambda_1)(D-\lambda_2)P^{-1} = 0$. The algebraic multiplicity of each eigenvalue equals its GM in this case.

(3) The matrix $A = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ has $p(\lambda) = \lambda(\lambda-1)$. The AM of each eigenvalue is 1. $(1,0),(1,1)$ form a basis of eigenvectors. (So the GM equals AM for each eigenvalue.) It is diagonalisable and $p(A) = 0$.

So we observe/comment/question :

(1) Not every matrix/linear map (even over finite-dimensional vector spaces) is diagonalisable. So which matrices are diagonalisable ? It seems that $AM = GM$ for all eigenvalues is necessary and sufficient. (Even if that is true, it is too much to check.)
(2) Nonetheless, it is possible that every matrix satisfies its own characteristic polynomial. If A is diagonalisable, then indeed, $p(A) = P(D-\lambda_1)(D-\lambda_2)\ldots P^{-1} = 0$. So we have proven the Cayley-Hamilton theorem for diagonalisable matrices.
(3) It appears that real symmetric matrices might be diagonalisable. (Explaining our change of variables in our ODE.)
(4) Non-symmetric matrices can also be diagonalisable.
(5) Can every matrix/linear map be brought to an upper triangular form at least ?