Consider a finite dimensional vector space V. Let $n = \dim(V)$. Many questions in mathematics and its applications often lead to a linear transformation $T : V \longrightarrow V$. The goal is to get a good understanding of T. For example, one could want to understand the iterations $T, T \circ T, T \circ T \circ T, \ldots$ of T. We will write T^k for the composition of T with itself k-times. For example, $T^4 = T \circ T \circ T \circ T$. It is convenient to write T^0 for the identity transformation.

Choosing a basis B of V allows us to identify T with its matrix ${}_BT_B$ (note that we use the same basis for both the domain V and the codomain V). The linear transformation T^k is then represented by the matrix $({}_BT_B)^k$, i.e. ${}_BT^k_B = ({}_BT_B)^k$. However, it is usually quite hard to understand high powers of a given matrix, unless the matrix has a particularly simple form. For example, it is easy to raise a diagonal matrix to any power:

$$\begin{bmatrix} a_1 & 0 & \ldots & 0 \\ 0 & a_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & a_n \end{bmatrix}^k = \begin{bmatrix} a_1^k & 0 & \ldots & 0 \\ 0 & a_2^k & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & a_n^k \end{bmatrix}$$

It is therefore natural to ask for a basis of V in which the matrix of T is as simple as possible. For example, can we always choose a basis in which the matrix of T is diagonal? Unfortunately, we will see that this is not possible in general, though there are many cases when it is possible. It is our goal to develop techniques which would allow us to answer such questions.

If D is another basis of V then the matrix ${}_DT_D$ of T in the basis D is given by

$${}_DT_D = {}_D I_B \ {}_BT_B \ {}_DI_B^{-1}$$

where ${}_DI_B$ is the transition matrix from basis B to basis D. Note that any invertible $n \times n$ matrix $M = [m_{i,j}]$ is equal to ${}_DI_B$ for some basis D of V. Indeed, let $M^{-1} = [k_{i,j}]$. If B consists of vectors $v_1, \ldots, v_n$ then let $w_j = k_{1,j}v_1 + k_{2,j}v_2 + \ldots + k_{n,j}v_n$ for $j = 1, \ldots, n$. The vectors $w_1, w_2, \ldots, w_n$ form a basis D of V and the transition matrix from the basis D to the basis B is M^{-1} (recall that the j-th column of ${}_BI_D$ consists of coordinates of w_j in the basis B). Thus $M = {}_D I_B$. It follows that finding a basis D in which the matrix of T is particularly simple is equivalent to finding an invertible matrix M such that $M({}_BT_B)M^{-1}$ is particularly simple. This motivates the following definition.

Definition. Two $n \times n$ matrices A_1, A_2 are called **similar** if $A_2 = MA_1M^{-1}$ for some invertible $n \times n$ matrix M.

Note that if A_1, A_2 are similar then A_2, A_1 are similar as well. Clearly any matrix is similar to itself and if A_1 is similar to A_2 and A_2 is similar to A_3 then A_1 is similar to A_3. Indeed, if $A_2 = MA_1M^{-1}$ and $A_3 = NA_2N^{-1}$ then $A_1 = (M^{-1})A_2(M^{-1})^{-1}$ and $A_3 = (NM)A_1(NM)^{-1}$.

Returning to our discussion of T, we see that a matrix A represents T in some basis if and only if it is similar to ${}_BT_B$. Our problem can be now stated as: among all matrices similar to a given matrix find one as simple as possible. Of course, this is not a very precise question as we do not specify what "simple" means. Getting a diagonal matrix would be ideal though (but, as we have already said, this is not always possible). We then introduce the following definition.

Definition. An $n \times n$ matrix A is called **diagonalizable** if it is simialr to a diagonal matrix, i.e. if MA_1M^{-1} is diagonal for some invertible $n \times n$ matrix M. A linear transformation $T : V \longrightarrow V$ is called **diagonalizable** if its matrix in some (any) basis is diagonalizable, i.e. if there is a basis of V in which the matrix of V is diagonal.

Suppose $w_1, ..., w_n$ is a basis D of V in which the first column of ${}_DT_D$ is of the form $(\lambda, 0, \ldots, 0)$. Recall that the first column provides coordinates of $T(w_1)$ in the basis D. This means that $T(w_1) = \lambda w_1$. Vectors with this property are therefore fundamental for the question of diagonalizability of T. This prompts the following definition.

Definition. Let $T : V \longrightarrow V$ be a linear transformation. A scalar λ is called an **eigenvalue** of T if there is a **non-zero** vector $v \in V$ such that $T(v) = \lambda v$. Any such v is called an **eigenvector** of T corresponding to eigenvalue λ.

If A is an $n \times n$ matrix then eigenvalues and eigenvectors of A are, by definition, the eigenvalues and eigenvectors of the linear transformation $L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$.

Note that λ is an eigenvalue of a matrix A if and only if the homogeneous system of equations $(A - \lambda I)x = 0$ has a non-zero solution. Recall that an $n \times n$ matrix is invertible iff it has rank n. We say that an $n \times n$ matrix is **singular** if it is not invertible. Thus

λ is an eigenvalue of an $n \times n$ matrix A if and only the rank of the matrix $A - \lambda I$ is smaller than n, i.e. if and only if $A - \lambda I$ is singular.

We have the following useful observation.

λ is an eigenvalue of the linear transformation $T : V \longrightarrow V$ if an only if λ is an eigenvalue for the matrix ${}_BT_B$ of T in some (any) basis of V. A vector $w \in V$ is an eigenvector of T corresponding to λ if and only if the vector of coordinates of w in the basis B is an eigenvector of the matrix ${}_BT_B$.

Indeed, if B consists of vectors $v_1, \ldots, v_n$ and if $w = t_1v_1 + \ldots t_nv_n$ then $T(w) = \lambda w$ if and only if ${}_BT_By = \lambda y$, where $y = (t_1, \ldots, t_n)$ is the vector of coordinates of w in the basis B.

In particular we have the following result.

Similar matrices have the same eigenvalues.

It is also easy to justify the last observation directly: if $Ax = \lambda x$ and M is invertible then $(MAM^{-1})(Mx) = MAx = M(Ax) = M(\lambda x) = \lambda(Mx)$. In other words, if x is an eigenvector of A corresponding to λ them Mx is an eigenvector of MAM^{-1} corresponding to λ.

We can state a criterion for a linear transformation to be diagonalizable.

A linear transformation $T : V \longrightarrow V$ is diagonalizable if an only if V has a basis consisting of eigenvectors of T.

Indeed, if $v_1, v_2, \ldots, v_n$ is a basis B of V consisting of eigenvectors, then there exist scalars λ_i such that $T(v_i) = \lambda_i v_i$, $i = 1, \ldots, n$. It follows that

$$
{}_BT_B = \begin{bmatrix} \lambda_1 & 0 \ldots & 0 \\ 0 & \lambda_2 \ldots & 0 \\ \vdots & \vdots \quad \vdots & \vdots \\ 0 & 0 \ldots & \lambda_n \end{bmatrix}
$$

is diagonal, so T is diagonalizable.

Conversely, if T is diagonalizable, then there is a basis D of V such that

$$_DT_D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

If D consists of vectors $w_1, \dots, w_n$ then we have $T(w_i) = \lambda_i w_i$ so D consists of eigenvalues of T.

The following simple observation is very useful.

Proposition. Let $T : V \longrightarrow V$ be a linear transformation and let λ be a scalar. The set

$$V(\lambda) = \{v \in V : T(v) = \lambda v\}$$

is a subspace of V called the **eigenspace of T corresponding to λ**.

Indeed, clearly $T(0) = 0 = \lambda \cdot 0$, so $0 \in V(\lambda)$. If $u, w \in V(\lambda)$ then $T(u) = \lambda u$ and $T(w) = \lambda w$. Thus

$$T(u + w) = T(u) + T(w) = \lambda u + \lambda w = \lambda(u + v)$$

and

$$T(au) = aT(u) = a(\lambda u) = \lambda(au)$$

for any scalar a. It follows that both $u + w$ and au are in $V(\lambda)$. This proves that $V(\lambda)$ is a subspace of V.

It is clear now that λ is an eigenvalue of T if and only if the eigenspace of T corresponding to λ contains a non-zero vector, i.e. if and only if $\dim(V(\lambda)) > 0$.

We end this note with the following important result.

Theorem. Let $\lambda_1, \dots, \lambda_s$ be pairwise distinct eigenvalues of a linear transformation T. Let u_i be an eigenvector corresponding to λ_i. The vectors $u_1, \dots, u_s$ are linearly independent.

Indeed, suppose they are not independent. Then there is $t < s$ such that $u_1, \dots, u_t$ are linearly independent and $u_{t+1} = a_1u_1 + \ldots + a_tu_t$ for some scalars a_t. Applying T we get $T(u_{t+1}) = T(a_1u_1 + \ldots + a_tu_t) = a_1T(u_1) + \ldots + a_tT(u_t)$. Since $T(u_i) = \lambda_i u_i$, we have

$$\lambda_{t+1}u_{t+1} = a_1\lambda_1u_1 + \ldots + a_t\lambda_tu_t.$$

Multiplying the equality $u_{t+1} = a_1u_1 + \ldots + a_tu_t$ by λ_{t+1} we get

$$\lambda_{t+1}u_{t+1} = a_1\lambda_{t+1}u_1 + \ldots + a_t\lambda_{t+1}u_t.$$

It follows that

$$0 = a_1(\lambda_1 - \lambda_{t+1})u_1 + \ldots + a_t(\lambda_t - \lambda_{t+1})u_t.$$

Since $u_1, \dots, u_t$ are linearly independent, we must have

$$0 = a_1(\lambda_1 - \lambda_{t+1}) = a_2(\lambda_2 - \lambda_{t+1}) = \ldots = a_t(\lambda_t - \lambda_{t+1}).$$

Since $\lambda_1, \dots, \lambda_s$ be pairwise distinct, this implies that $a_1 = \ldots = a_t = 0$. However this means that $u_{t+1} = a_1u_1 + \ldots + a_tu_t = 0$, which is false. We see that our assumption that $u_1, \dots, u_s$ are linearly dependent leads to a contradiction. Thus $u_1, \dots, u_s$ are linearly independent

The last theorem has several important corollaries.

Corollary. Let $\lambda_1, \ldots, \lambda_s$ be pairwise distinct eigenvalues of a linear transformation T. Let $w_{i,1}, w_{i,2}, \ldots, w_{i,d_i}$ be a basis of $V(\lambda_i)$, $i = 1, \ldots, s$. Then the sequence

$$w_{1,1}, \ldots, w_{1,d_1}, w_{2,1}, \ldots, w_{2,d_2}, \ldots, w_{s,1}, \ldots, w_{s,d_s}$$

is linearly independent.

Indeed suppose that

$$a_{1,1}w_{1,1} + \ldots + a_{1,d_1}w_{1,d_1} + a_{2,1}w_{2,1} + \ldots + a_{2,d_2}w_{2,d_2} + \ldots + +a_{s,1}w_{s,1} + \ldots + a_{s,d_s}w_{s,d_s} = 0.$$

Let $u_i = a_{i,1}w_{i,1} + \ldots + a_{i,d_i}w_{i,d_i}$. Thus $u_1 + u_2 + \ldots + u_s = 0$. Note that $u_i \in V(\lambda_i)$. By our last theorem, the non-zero vectors among $u_1, \ldots u_s$ are linearly independent. This means that all u_i must be actually zero. Since $w_{i,1}, w_{i,2}, \ldots, w_{i,d_i}$ is a basis of $V(\lambda_i)$, we have $a_{i,1} = \ldots = a_{i,d_i} = 0$.

Corollary. Let $\lambda_1, \ldots, \lambda_s$ be pairwise distinct eigenvalues of a linear transformation T. Then

$$\dim(V(\lambda_1)) + \ldots + \dim(V(\lambda_s)) \leq \dim(V).$$

Corollary. Let $T : V \longrightarrow V$ be a linear transformation. Then T has at most $\dim(V)$ different eigenvalues. In particular, an $n \times n$ matrix has at most n distinct eigenvalues.

Corollary. Let $T : V \longrightarrow V$ be a linear transformation. Let $\lambda_1, \ldots, \lambda_s$ be all the eigenvalues of T. Then T is diagonalizable if and only if

$$\dim(V(\lambda_1)) + \ldots + \dim(V(\lambda_s)) = \dim(V).$$

In particular, if T has $\dim(V)$ distinct eigenvalues, then T is diagonalizable.

Example 1. Consider the space $\mathbb{P}_3$ of polynomials of degree at most 3. Let $T : \mathbb{P}_3 \longrightarrow \mathbb{P}_3$ be the differentiation: $T(f) = f'$. λ is an eigenvalue of T if $f' = \lambda f$ for some non-zero polynomial f. If $\lambda \neq 0$ then f' has degree lower than λf, hence we can not have $f' = \lambda f$. If $\lambda = 0$ then $f' = 0$, so f is constant. We see that 0 is the only eigenvalue of T and the corresponding eigenspace has dimension 1 (it consists of constant polynomials). Thus T is not diagonalizable. Since $T(1) = 0$, $T(x) = 1$, $T(x^2) = 2x$ and $T(x^3) = 3x^2$, the matrix of T in the basis $1, x, x^2, x^3$ is

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Thus this matrix is not diagonalizable.

Example 2. Consider the matrix

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

and let $T = L_A : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be the corresponding linear transformation. Thus A is the matrix of T in the standard basis. We claim that 3 and 5 are eigenvalues of T. Indeed,

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

We see that $A-3I$ is singular, so 3 is indeed an eigenvalue of A. The corresponding eigenspace consists of solution to $(A-3I)x=0$, i.e.

$$\begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The reduced row echelon form of $\begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 1 & 0 & 1 \end{bmatrix}$ is $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Thus space of solutions is two dimensional (we have 2 free variables x_2 and x_3) and a basis of solutions is $(0,1,0),(-1,0,1)$.

Similarly, we have

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

We see that $A-3I$ is singular, so 5 is also an eigenvalue of A. The corresponding eigenspace consists of solution to $(A-5I)x=0$, i.e.

$$\begin{bmatrix} -1 & 0 & 1 \\ 2 & -2 & 2 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The reduced row echelon form of $\begin{bmatrix} -1 & 0 & 1 \\ 2 & -2 & 2 \\ 1 & 0 & -1 \end{bmatrix}$ is $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$. Thus space of solutions is one dimensional (we have 1 free variable x_3) and a basis of solutions is $(1,2,1)$.

According our theorem, the vectors $(0,1,0),(-1,0,1),(1,2,1)$ are linearly independent. Thus they form a basis B of $\mathbb{R}^3$ which consists of eigenvectors. Thus T is diagonalizable. The transition matrix ${}_EI_B$ from the basis B to the standard basis E is $\begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix}$. The transition matrix ${}_BI_E$ from the standard basis to the basis B is the inverse of $\begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix}$. A simple computation yields $\begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & 1 & -1 \\ -1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \end{bmatrix}$. Thus

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

Thus A is diagonalizable and it is similar to the diagonal matrix $\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$.

Example 3. Suppose that we have a particle which can be at one of two different states 1 and 2. Every second the particle in state i can change to the other state j with probability $p_{i,j}$ or it can stay at the state i with probability $p_{i,i} = 1 - p_{i,j}$. We would like to know the probability $p_{i,j}(n)$ that a particle originally in state i is in state j after n seconds. It is a nice exercise to see that the answer is given in terms of the 2×2 matrix $P = [p_{i,j}]$. Namely $p_{i,j}(n)$ is simply the (i,j) entry of the matrix P^n.

To make things more concrete, suppose that $p_{1,1} = 1/3$ and $p_{2,2} = 1/2$ so that the matrix P is $\begin{bmatrix} 1/3 & 2/3 \\ 1/2 & 1/2 \end{bmatrix}$. We consider P as the matrix of a linear transformation $T = L_P : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ in the standard basis E of $\mathbb{R}^2$. λ is an eigenvalue of P if and only if the matrix $\begin{bmatrix} \frac{1}{3}-\lambda & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2}-\lambda \end{bmatrix}$ is singular. Performing on this matrix the row operations $S_{1,2}$, $D_1(2)$, $D_2(3)$, and $E_{2,1}(3\lambda-1)$ we get

$$\begin{bmatrix} 1 & 1-2\lambda \\ 0 & 2+(3\lambda-1)(1-2\lambda) \end{bmatrix}.$$

The last matrix is singular if and only if $2+(3\lambda-1)(1-2\lambda)=0$, i.e. $6\lambda^2-5\lambda-1=0$. The solutions to this quadratic equations are $\lambda=1$ and $\lambda=-1/6$. Thus 1 and $-1/6$ are eigenvalues of the matrix P. We now easily find that $v_1=(1,1)$ is an eigenvector corresponding to the eigenvalue 1 (so $T(1,1)=(1,1)$) and $v_2=(-4,3)$ is an eigenvector corresponding to the eigenvalue $-1/6$ (so $T(-4,3)=\frac{-1}{6}(-4,3)$). It follows that v_1,v_2 is a basis B of $\mathbb{R}^2$ consisting of eigenvectors of T. Thus P is diagonalizable: $CPC^{-1}=\begin{bmatrix} 1 & 0 \\ 0 & -1/6 \end{bmatrix}$, where C is the transition matrix from the standard basis E to the basis B. Now it is easy to get C^{-1}, which is the transition matrix from B to E, so $C^{-1}=\begin{bmatrix} 1 & -4 \\ 1 & 3 \end{bmatrix}$. Inverting this matrix, we get

$$C=\begin{bmatrix} \frac{3}{7} & \frac{4}{7} \\ \frac{-1}{7} & \frac{1}{7} \end{bmatrix}.$$

Note that for any matrix A, any invertible matrix M, and any positive integer k we have $(M^{-1}AM)^k = M^{-1}A^kM$. Since $P=C^{-1}\begin{bmatrix} 1 & 0 \\ 0 & -1/6 \end{bmatrix}C$, we get

$$P^k=C^{-1}\begin{bmatrix} 1 & 0 \\ 0 & \frac{-1}{6} \end{bmatrix}^k C=\begin{bmatrix} 1 & -4 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & (\frac{-1}{6})^k \end{bmatrix}\begin{bmatrix} \frac{3}{7} & \frac{4}{7} \\ \frac{-1}{7} & \frac{1}{7} \end{bmatrix}=\begin{bmatrix} \frac{3}{7}+\frac{4}{7}(\frac{-1}{6})^k & \frac{4}{7}-\frac{4}{7}(\frac{-1}{6})^k \\ \frac{3}{7}-\frac{3}{7}(\frac{-1}{6})^k & \frac{4}{7}+\frac{3}{7}(\frac{-1}{6})^k \end{bmatrix}$$

As a consequence of our discussion, note that for k very large the probability that the particle ends up in state 1 practically does not depend on where it started and is about 3/7. The probability that the particle ends up in state 2 is about 4/7.

Exercise. Solve the above problem in general.
Hint. The cases when both $p_{1,1}$ and $p_{2,2}$ are 0 or both are 1 requires separate treatment, but is easy. All other cases are handled by the same method as in our example. Just note that the vectors $v_1=(1,1)$ and $v_2=(-p_{1,2},p_{2,1})$ are eigenvectors of P.