As motivation, consider this list of problems that have arisen naturally in our theoretical study of linear algebra:

1. Determine whether or not a given list of vectors spans a vector space or subspace.
2. Determine whether or not a given list of vectors is linearly independent.
3. Given a list of linearly dependent vectors which span a space, reduce them to a basis of that space.
4. Given a vector space, or subspace, find its dimension and find a basis for that space.
5. Given a basis, find the coordinates of some vector relative to that basis.
6. Solve systems of linear equations (related to the previous item).
7. Given coordinates of one vector relative to a basis, compute the coordinates of that vector relative to some other basis.
8. Given a basis, convert it to an orthonormal basis (Gram-Schmidt).
9. Given a subspace and a vector, compute the projection of that vector onto the subspace.
10. Classify all possible inner product formulas on a vector space.

So far we have looked at special cases of these problems where we can compute the answers conceptually or theoretically or with the aid of pictures. But if there are lots of complicated numbers to work with, or if the dimension is too high to draw a picture, the concepts don't work by themselves.

However, there is a tool which allows us to rethink all the problems posed above (and others) in a fashion that makes the answers to all the questions above "computable" by some kind of "recipe". The tool is matrix theory.

Definition 1 *Given a field F and positive integers m and n, an $m \times n$* matrix *with entries in F is an array of numbers $a_{ij} \in F$ where $1 \leq i \leq m$ and $1 \leq j \leq n$. We denote matrices by capital letters (usually); a matrix with entries a_{ij} is usually denoted A. We arrange the entries of the matrix in a rectangle as follows:*

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & \cdots & a_{mn} \end{pmatrix}$$

In particular, a_{ij} is the entry of A in the i^{th} row and j^{th} column. m is the number of rows of A; n is the number of columns of A.

The set of matrices of size $m \times n$ with entries in F is denoted $M_{mn}(F)$. If $m = n$, the matrix is called *square*; the set of $n \times n$ square matrices with entries in F is denoted $M_n(F)$. The *diagonal entries* of a $m \times n$ matrix are the numbers $a_{11}, a_{22}, a_{33}, ..., a_{nn}$ (this list may stop at a_{mm} depending on when you run out of entries). A matrix is called *diagonal* if it is square and all its nondiagonal entries are zero. The *trace* of a matrix is the sum of the diagonal entries of that matrix. The $n \times n$ *identity matrix*, denoted I or I_n, is the diagonal $n \times n$ matrix with all diagonal entries equal to 1.

Two matrices are *equal* if they are the same size and if all their entries coincide, i.e. $A = B$ if they are both $m \times n$ and if $a_{ij} = b_{ij}$ for all i, j.

Vector space operations: Given two matrices $A, B \in M_{mn}(F)$ and a scalar $r \in F$, we define $A + B \in M_{mn}(F)$ by

$$A + B = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} & \cdots & a_{2n}+b_{2n} \\ a_{31}+b_{11} & a_{32}+b_{11} & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & \cdots & a_{mn}+b_{mn} \end{pmatrix};$$

equivalently we set $(a+b)_{ij} = a_{ij} + b_{ij}$ for all i, j. We define $rA \in M_{mn}(F)$ by

$$rA = \begin{pmatrix} ra_{11} & ra_{12} & ra_{13} & \cdots & ra_{1n} \\ ra_{21} & ra_{22} & ra_{23} & \cdots & ra_{2n} \\ ra_{31} & ra_{32} & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ ra_{m1} & ra_{m2} & \cdots & \cdots & ra_{mn} \end{pmatrix};$$

equivalently we can define this by setting $(ra)_{ij} = r(a_{ij})$ for all i, j. These two operations make $M_{mn}(F)$ into a vector space over F; the additive identity element is the $m \times n$ *zero matrix*

$$0 = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

The dimension of $M_{mn}(F)$ is mn; a basis for $M_{mn}(F)$ is

$$\{\Delta_{11}, \Delta_{12}, \Delta_{13}, ..., \Delta_{1n}, \Delta_{21}, \Delta_{22}, ..., \Delta_{2n}, \Delta_{31}, ..., \Delta_{m1}, ..., \Delta_{mn}\}$$

where Δ_{ij} is the $m \times n$ matrix with all entries equal to zero except the $(i,j)-$entry, which is 1. As a special case of this, we see that $\dim M_n(F) = n^2$.

Transposes, Hermitians, and conjugates:

Definition 2 *Given $A \in M_{mn}(F)$, define the* transpose *of A, denoted A^T or A^t, to be the $n \times m$ matrix satisfying $(a^T)_{ij} = a_{ji}$ for all i, j.*

Definition 3 *Given $A \in M_{mn}(\mathbb{C})$, define the* conjugate *of A, denoted $\overline{A}$, to be the $m \times n$ matrix satisfying $(\overline{a})_{ij} = \overline{a_{ij}}$ for all i, j.*

Definition 4 *Given $A \in M_{mn}(\mathbb{C})$, define the* Hermitian *of A, denoted A^H or A^h and called "A Hermitian", to be the $n \times m$ matrix $A^H = \overline{A^T}$.* Note: *the "Hermitian of a matrix" is different from "a Hermitian matrix". The first usage of Hermitian (as defined here) is as a noun, the second usage (which will be defined later) is as an adjective.*

Associating vectors in F^n to column vectors: We associate vectors in F^n to $n \times 1$ matrices as follows:

$$\mathbf{x} = (x_1, x_2, ..., x_n) \in F^n \qquad \leftrightarrow \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in M_{n1}(F).$$

In particular, an $n \times 1$ matrix is also called a *column vector*. A column vector is the same thing as a vector in F^n. Given a vector $\mathbf{x} = (x_1, ..., x_n) \in F^n$, if we want to think of that vector as a row vector, we take the transpose of $\mathbf{x}$:

$$\mathbf{x}^T = \begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_n \end{pmatrix}.$$

Rows and columns of a matrix: Given an $m \times n$ matrix A as above, the vectors

$$\begin{gathered}(a_{11}, a_{12}, a_{13}, ..., a_{1n}),\\ (a_{21}, a_{22}, a_{23}, ..., a_{2n}),\\ (a_{31}, a_{32}, ..., a_{3n}),\\ ...,\\ (a_{m1}, a_{m2}, ..., a_{mn})\end{gathered}$$

are called the *rows* of A; note that each row of A is an element of F^n. The vectors

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, ..., \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

are called the *columns* of A; note that each column of A is an element of F^m.

Matrix multiplication:

Definition 5 *Given matrices $A \in M_{mn}(F)$ and $B \in M_{pq}(F)$, if $n = p$ then we can define the* product AB, *which is an $m \times q$ matrix AB defined by*

$$(ab)_{ij} = \sum_{k=1}^{n(=p)} a_{ik} b_{kj}.$$

(If $n \neq p$, AB is undefined.) If A is a square matrix, we write A^2 for AA, A^3 for AAA, etc.

In general matrix multiplication is not commutative: $AB \neq BA$ most of the time, even if both products are defined.

Properties of elementary matrix operations: Let A, B, C be matrices with entries in the same field F, let I be an identity matrix of the appropriate size and let $r \in F$. Then, **so long as everything is defined**, we have:

1. $IA = A$ and $BI = B$.
2. $A(BC) = (AB)C$;
3. $k(AB) = (kA)B = A(kB)$;
4. $A(B + C) = AB + AC$;
5. $(A + B)C = AC + AC$;
6. $(A^T)^T = A$ and $(A^H)^H = A$;
7. $(A^T)^H = (A^H)^T = \overline{A}$;
8. $tr(A^T) = tr(A)$;
9. $tr(A^H) = tr(\overline{A}) = \overline{tr(A)}$;
10. $(rA)^T = rA^T$ and $(rA)^H = \overline{r}A^H$;
11. $(A + B)^T = A^T + B^T$ and $(A + B)^H = A^H + B^H$;
12. $(AC)^T = C^T A^T$ and $(AC)^H = C^H A^H$;
13. $tr(A + B) = trA + trB$;
14. $tr(AD) = tr(DA)$;