Math 115A - Week 6
Textbook sections: 3.1-5.1
Topics covered:

- Review: Row and column operations on matrices
- Review: Rank of a matrix
- Review: Inverting a matrix via Gaussian elimination
- Review: Determinants

* * * * *

Review: Row and column operations on matrices

- We now quickly review some material from Math 33A which we will need later in this course. The first concept we will need is that of an *elementary row operation.*

- An *elementary row operation* takes an $m \times n$ matrix as input and returns a different $m \times n$ matrix as output. (In other words, each elementary row operation is a map from $M_{m\times n}(\mathbf{R})$ to $M_{m\times n}(\mathbf{R})$. There are three types of elementary row operations:

- **Type 1 (row interchange)**. This type of row operation interchanges row i with row j for some $i, j \in \{1, 2, \ldots, m\}$. For instance, the operation of interchanging rows 2 and 4 in a 4×3 matrix would change

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix}$$

to

$$\begin{pmatrix} 1 & 2 & 3 \\ 10 & 11 & 12 \\ 7 & 8 & 9 \\ 4 & 5 & 6 \end{pmatrix}.$$

Observe that the final matrix can be obtained from the initial matrix by multiplying on the left by the 4×4 matrix

$$E := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

which is the identity matrix with rows 2 and 4 switched. (Why?) Thus for instance

$$\begin{pmatrix} 1 & 2 & 3 \\ 10 & 11 & 12 \\ 7 & 8 & 9 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix}$$

and more generally, if A is a 4×3 matrix, then the interchange of rows 2 and 4 replaces A with EA. We refer to E as an 4×4 *elementary matrix of type 1.*

- Also observe that row interchange is its own inverse; if one replaces A with EA, and then replaces EA with EEA (i.e. we interchange rows 2 and 4 twice), we get back to where we started, because $EE = I_4$.

- **Type 2 (row multiplication)** This type of elementary row operation takes a row i and multiplies it by a *non-zero* scalar c. For instance, the elementary row operation that multiplies row 2 by 10 would map

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix}$$

to

$$\begin{pmatrix} 1 & 2 & 3 \\ 40 & 50 & 60 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix}.$$

This operation is the same as multiplying on the left by the matrix

$$E := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

which is what one gets by starting with the identity matrix and multiplying row 2 by 10. (Why?) We call E an example of a 4×4 *elementary matrix of type 2.*

- Row multiplication is invertible; the operation of multiplying a row i by a non-zero scalar c is inverted by multiplying a row i by the non-zero scalar $1/c$. In the above example, the inverse operation is given by

$$E^{-1} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/10 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

i.e. to invert the operation of multiplying row 2 by 10, we then multiply row 2 by 1/10.

- **Type 3 (row addition)** For this row operation we need two rows i, j, and a scalar c. The row operation adds c multiples of row i to row j. For instance, if one were to add 10 multiples of row 2 to row 3, then

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix}$$

would become

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 47 & 58 & 69 \\ 10 & 11 & 12 \end{pmatrix}.$$

Equivalently, this row operation amounts to multiplying the original matrix on the left by the matrix

$$E := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 10 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

which is what one gets by starting with the identity matrix and adding 10 copies of row 2 to row 3. (Why?) We call E an example of a 4×4 *elementary matrix of type* 3. It has an inverse

$$E^{-1} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -10 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix};$$

thus to invert the operation of adding 10 copies of row 2 to row 3, we subtract 10 copies of row 2 to row 3 instead.

- Thus to summarize: there are special matrices, known as *elementary row matrices*, and an elementary row operation amounts to multiplying a given matrix on the left by one of these elementary row matrices. Each of the elementary row matrices is invertible, and the inverse of an elementary matrix is another elementary matrix. (In the above discussion, we only verified this for specific examples of elementary matrices, but it is easy to see that the same is true for general elementary matrices. See the textbook).

- There are also *elementary column operations*, which are very similar to elementary row operations, but arise from multiplying a matrix on the *right* by an elementary matrix, instead of the left. For instance, if one multiplies a matrix A with 4 columns on the right by the elementary matrix

$$E := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

then this amounts to switching column 2 and column 4 of A (why?). This is a type 1 (column interchange) elementary column move. Similarly, if one multiplies A on the right by

$$E := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

then this amounts to multiplying column 2 of A by 10 (why?). This is a type 2 (column multiplication) elementary column move. Finally, if one multiplies A on the right by

$$E := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 10 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

then this amounts to adding 10 copies of column 3 to column 2 (why?). This is a type 3 (column addition) elementary column move.

- Elementary row (or column) operations have several uses. One important use is to simplify systems of linear equations of the form $Ax = b$, where A is some matrix and x, b are vectors. If E is an elementary matrix, then the equation $Ax = b$ is equivalent to $EAx = Eb$ (why are these two equations equivalent? Hint: E is invertible). Thus one can simplify the equation $Ax = b$ by performing the same row operation to both A and b simultaneously (one can concatenate A and b into a single (artificial) matrix in order to do this). Eventually one can use row operations to reduce A to row-echelon form (in which each row is either zero, or begins with a 1, and below each such 1 there are only zeroes), at which point it becomes straightforward to solve for $Ax = b$ (or to determine that there are no solutions). We will not review this procedure here, because it will not be necessary for this course; see however the textbook (or your Math 33A notes) for more information. However, we will remark that every matrix can be row-reduced into row-echelon form (though there is usually more than one way to do so).

- Another purpose of elementary row or column operations is to determine the *rank* of a matrix, which is a more precise measurement of its invertibility. This will be the purpose of the next section.

* * * * *

Rank of a matrix

- Recall that the *rank* of a linear transformation $T : V \to W$ is the dimension of its range $R(T)$. Rank has a number of uses, for instance it can be used to tell whether a linear transformation is invertible:

- **Lemma.** Let V and W be n-dimensional spaces, and let $\dim(V) = \dim(W) = n$. Let $T : V \to W$ be a linear transformation. Then T is invertible if and only if $\text{rank}(T) = n$ (i.e. T has the maximum rank possible).

- **Proof.** If $\text{rank}(T) = n$, then $R(T)$ has the same dimension n as W. But $R(T)$ is a subspace of W, so this forces $R(T)$ to actually equal W (see Theorem 2 of Week 2 notes). Thus T is onto. Also, from the dimension theorem we see that $\text{nullity}(T) = 0$, and so T is one-to-one. Thus T is invertible.

- Conversely, if T is invertible, then it is one-to-one, hence $\text{nullity}(T) = 0$, and hence by the dimension theorem $\text{rank}(T) = n$. $\square$

- One interesting thing about rank is that if you multiply a linear transformation on the left or right by an invertible transformation, then the rank doesn't change:

- **Lemma 1.** Let $T : V \to W$ be a linear transformation from one finite-dimensional space to another, let $S : U \to V$ be an invertible transformation, and let $Q : W \to Z$ be another invertible transformation. Then

$$\text{rank}(T) = \text{rank}(QT) = \text{rank}(TS) = \text{rank}(QTS).$$

- **Proof.** First let us show that $\text{rank}(T) = \text{rank}(TS)$. To show this, we first compute the ranges of $T : V \to W$ and $TS : U \to W$. By definition of range, $R(T) = T(V)$, the image of V under T. Similarly,

$R(TS) = TS(U)$. But since $S : U \to V$ is invertible, it is onto, and so $S(U) = V$. Thus

$$R(TS) = TS(U) = T(S(U)) = T(V) = R(T)$$

and so

$$\text{rank}(TS) = \text{rank}(T).$$

A similar argument gives that $\text{rank}(QTS) = \text{rank}(QT)$ (just replace T by RT in the above. To finish the argument we need to show that $\text{rank}(QT) = \text{rank}(T)$. We compute ranges again:

$$R(QT) = QT(V) = Q(T(V)) = Q(R(T)),$$

so that

$$\text{rank}(QT) = \dim(Q(R(T))).$$

But since Q is invertible, we have

$$\dim(Q(R(T))) = \dim(R(T)) = \text{rank}(T)$$

(see Q3 of the midterm!). Thus $\text{rank}(QT) = \text{rank}(T)$ as desired. $\square$.

- Now to compute the rank of an arbitrary linear transformation can get messy. The best way to do this is to convert the linear transform to a matrix, and compute the rank of that.

- **Definition** If A is a matrix, then the *rank* of A, denoted $\text{rank}(A)$, is defined by $\text{rank}(A) = \text{rank}(L_A)$.

- **Example** Consider the matrix

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

Then L_A is the transformation

$$L_A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 0 \end{pmatrix}.$$

The range of this operator is thus three-dimensional (why?) and so the rank of A is 3.

- Let A be an $m \times n$ matrix, so that L_A maps $\mathbf{R}^n$ to $\mathbf{R}^m$. Let $(e_1, e_2, \ldots, e_n)$ be the standard basis for $\mathbf{R}^n$. Since $e_1, \ldots, e_n$ span $\mathbf{R}^n$, we see that $L_A(e_1), L_A(e_2), \ldots, L_A(e_n)$ span $R(L_A)$ (see Theorem 3 of Week 3 notes). Thus the rank of A is the dimension of the space spanned by $L_A(e_1), L_A(e_2), \ldots, L_A(e_n)$. But $L_A(e_1)$ is simply the first column of A (why?), $L_A(e_2)$ is the second column of A, etc. Thus we have shown

- **Lemma 2.** The rank of a matrix A is equal to the dimension of the space spanned by its columns.

- **Example** If A is the matrix used in the previous example, then the rank of A is the dimension of the span of the columns $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$; this span is easily seen to be three-dimensional.

- As one corollary of this Lemma, if only k of the columns of a matrix are non-zero, then the rank of the matrix can be at most k (though it could be smaller than k; can you think of an example?).

- This Lemma does not necessarily make computing the rank easy, because finding the dimension of the space spanned by the columns could be difficult. However, one can use elementary row or column operations to simplify things. From Lemma 1 we see that if E is an elementary matrix and A is a matrix, then EA or AE has the same rank as A (provided that the matrix multiplication makes sense, of course). Thus elementary row or column operations do not change the rank. Thus, we can use these operations to simplify the matrix into row-echelon form.

- **Lemma 3.** Let A be a matrix in row-echelon form (thus every row is either zero, or begins with a 1, and each of those 1s has nothing but 0s below it). Then the rank of A is equal to the number of non-zero rows.

- **Proof.** Let's say that A has k non-zero rows, so that we have to show that A has rank k. Each column of A can only has the top k entries non-zero; all entries below that must be zero. Thus the span of the columns of A is contained in the k-dimensional space

$$V = \{ \begin{pmatrix} x_1 \\ \ldots \\ x_k \\ 0 \\ \ldots \\ 0 \end{pmatrix} : x_1, \ldots, x_k \in \mathbf{R} \},$$

and so the rank is at most k.

- Now we have to show that the rank is at least k. To do this it will suffice to show that every vector in V is in the span of the columns of A, since this will mean that the span of the columns of A is exactly the k-dimensional space V. So, let us pick a vector $v := \begin{pmatrix} x_1 \\ \ldots \\ x_k \\ 0 \\ \ldots \\ 0 \end{pmatrix}$ in V. Since A is in row-echelon form, the k^{th} row of A must contain a 1 somewhere, which means that there is a column whose k^{th} entry is 1 (with all entries below that equal to 0). If we subtract x_k multiples of this column from v, then we get a new vector whose k^{th} entry (and all the ones below it) are zero.

- Now we look at the $(k-1)^{th}$ row. Again, since we are in row-echelon form, there is a 1 somewhere, with 0s below it; this implies that there is a column whose $(k-1)^{th}$ entry is 1 (with all entries below that equal to 0). Thus we can subtract a multiple of this vector to get a new vector whose $(k-1)^{th}$ entry (and all the ones below it) are zero.

- Continuing in this fashion we can subtract multiples of various columns of A from v until we manage to zero out all the entries. In other words, we have expressed v as a linear combination of columns in A, and hence

v is in the span of the columns. Thus the span of the columns is exactly V, and we are done. $\square$.

- Thus we now have a procedure to compute the rank of a matrix: we row reduce (or column reduce) until we reach row-echelon form, and then just count the number of non-zero rows. (Actually, one doesn't necessarily have to reduce all the way to row-echelon form; it may be that the rank becomes obvious some time before then, because the span of the columns can be determined by inspection).

- If one only uses elementary row operations, then usually one cannot hope to make the matrix much simpler than row-echelon form. But if one is allowed to also use elementary column operations, then one can get the matrix into a particularly simple form:

- **Theorem 4.** Let A be an $m \times n$ matrix with rank r. Then one can use elementary row and column matrices to place A in the form

$$\begin{pmatrix} I_r & 0_{r \times n-r} \\ 0_{m-r \times r} & 0_{m-r \times n-r} \end{pmatrix},$$

  where I_r is the $r \times r$ identity matrix, and $0_{m \times n}$ is the $m \times n$ zero matrix. (Thus, we have reduced the matrix to nothing but a string of r ones down the diagonal, with everything else being zero.

- **Proof** To begin with, we know that we can use elementary row operations to place A in row-echelon form. Thus the first r rows begin with a 1, with 0s below the 1, while the remaining $m - r$ rows are entirely zero.

- Now consider the first row of this reduced matrix; let's suppose that it is not identically zero. After some zeroes, it has a 1, and then some other entries which may or may not be zero. But by subtracting multiples of the column with 1 in it from the columns with other non-zero entries (i.e. type 3 column operations), we can make all the other entries in this row equal to zero. Note that these elementary column operations only affect the top row, leaving the other rows unchanged, because the column with 1 in it has 0s everywhere else.

- A similar argument then allows one to take the second row of the reduced matrix and make all the entries (apart from the leading 1) equal to 0. And so on and so forth. At the end of this procedure, we get that the first r rows each contain one 1 and everything else being zero. Furthermore, these 1s have 0s both above and below, so they lie in different columns. Thus by switching columns appropriately (using type 1 column operations) we can get into the form required for the Theorem. $\square$

- Let A be an $m \times n$ matrix with rank r. Every elementary row operation corresponds to multiplying A on the left by an $m \times m$ elementary matrix, while every elementary column operation corresponds to multiplying A on the right by an $n \times n$ elementary matrix. Thus by Theorem 4, we have

$$E_1 E_2 \dots E_a A F_1 F_2 \dots F_b = \begin{pmatrix} I_r & 0_{r \times n-r} \\ 0_{m-r \times r} & 0_{m-r \times n-r} \end{pmatrix},$$

where $E_1, \dots, E_a$ are elementary $m \times m$ matrices, and $F_1 F_2 \dots F_b$ are elementary $n \times n$ matrices. All the elementary matrices are invertible. After some matrix algebra, this becomes

$$A = E_a^{-1} \dots E_2^{-1} E_1^{-1} \begin{pmatrix} I_r & 0_{r \times n-r} \\ 0_{m-r \times r} & 0_{m-r \times n-r} \end{pmatrix} F_b^{-1} \dots F_2^{-1} F_1^{-1}$$

(why did the order of the matrices get reversed?). We have thus proven

- **Proposition 5.** Let A be an $m \times n$ matrix with rank r. Then we have an $m \times m$ matrix B which is a product of elementary matrices and an $n \times n$ matrix C, also a product of elementary matrices such that

$$A = B \begin{pmatrix} I_r & 0_{r \times n-r} \\ 0_{m-r \times r} & 0_{m-r \times n-r} \end{pmatrix} C.$$

- Note (from Q6 of Assignment 4) that the product of invertible matrices is always invertible. Thus the matrices B and C above are invertible.

- Proposition 5 is an example of a *factorization theorem*, which takes a general matrix and factors it into simpler pieces. There are many other examples of factorization theorems which you will encounter later in the 115 sequence, and they have many applications.

- Some more properties of rank. We know from Lemma 1 that rank of a *linear transformation* is unchanged by multiplying on the left or right by invertible transformations. Given the close relationship between linear transformations and matrices, it is unsurprising that the same thing is true for matrices:

- **Lemma 6.** Let A be an $m \times n$ matrix, B be an $m \times m$ invertible matrix, and C be an $n \times n$ invertible matrix. Then

$$\mathrm{rank}(A) = \mathrm{rank}(BA) = \mathrm{rank}(AC) = \mathrm{rank}(BAC).$$

- **Proof.** Since B is invertible, so is L_B (see Theorem 10 from Week 4 notes). Similarly L_C is invertible. From Lemma 1 we have

$$\mathrm{rank}(L_A) = \mathrm{rank}(L_B L_A) = \mathrm{rank}(L_A L_C) = \mathrm{rank}(L_B L_A L_C).$$

Since $L_B L_A = L_{BA}$, etc. we thus have

$$\mathrm{rank}(L_A) = \mathrm{rank}(L_{BA}) = \mathrm{rank}(L_{AC}) = \mathrm{rank}(L_{BAC}).$$

The claim then follows from the definition of rank for matrices. $\square$

- Note that Lemma 6 is consistent with Proposition 5, since the matrix $\begin{pmatrix} I_r & 0_{r \times n-r} \\ 0_{m-r \times r} & 0_{m-r \times n-r} \end{pmatrix}$ has rank r.

- One important consequence of the above theory concerns the rank of a *transpose* A^t of a matrix A. Recall that the transpose of an $m \times n$ matrix is the $n \times m$ matrix obtained by reflecting around the diagonal, so for instance

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{pmatrix}^t = \begin{pmatrix} 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \end{pmatrix}.$$

Thus transposes swap rows and columns. From the definition of matrix multiplication it is easy to verify the identity $(AB)^t = B^t A^t$ (why?). In particular, if A is invertible, then $I = I^t = (AA^{-1})^t = A^t (A^{-1})^t$, which implies that A^t is also invertible.

- **Lemma 7** Let A be an $m \times n$ matrix with rank r. Then A^t has the same rank as A.

- From Proposition 5 we have

$$A = B \begin{pmatrix} I_r & 0_{r \times n-r} \\ 0_{m-r \times r} & 0_{m-r \times n-r} \end{pmatrix} C.$$

  Taking transposes of both sides we obtain

$$A^t = C^t \begin{pmatrix} I_r & 0_{r \times m-r} \\ 0_{n-r \times r} & 0_{n-r \times m-r} \end{pmatrix} B^t.$$

  The inner matrix on the right-hand side has rank r. Since B and C are invertible, so are B^t and C^t, and so by Lemma 6 A^t has rank r, and we are done. $\square$

- From Lemma 7 and Lemma 2 we thus have

- **Corollary 8.** The rank of a matrix is equal to the dimension of the space spanned by its *rows*.

- As one corollary of this Lemma, if only k of the rows of a matrix are non-zero, then the rank of the matrix can be at most k.

- Finally, we give a way to compute the rank of any linear transformation from one finite-dimensional space to another.

- **Lemma 9.** Let $T : V \to W$ be a linear transformation, and let β and γ be finite bases for V and W respectively. Then $\mathrm{rank}(T) = \mathrm{rank}([T]_\beta^\gamma)$.

- **Proof.** Suppose V is n-dimensional and W is m-dimensional. Then the co-ordinate map $\phi_\beta : V \to \mathbf{R}^n$ defined by $\phi_\beta(v) := [v]^\beta$ is an isomorphism, as is the co-ordinate map $\phi_\gamma : W \to \mathbf{R}^m$ defined by $\phi_\gamma(w) := [w]^\gamma$. Meanwhile, the map $L_{[T]_\beta^\gamma}$ is a linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^m$. The identity

$$[Tv]^\gamma = [T]_\beta^\gamma [v]^\beta$$

  can thus be rewritten as

$$\phi_\gamma(Tv) = L_{[T]_\beta^\gamma} \phi_\beta(v)$$

and thus

$$\phi_\gamma T = L_{[T]_\beta^\gamma} \phi_\beta$$

and hence (since ϕ_β is invertible)

$$T = \phi_\gamma^{-1} L_{[T]_\beta^\gamma} \phi_\beta.$$

Taking rank of both sides and using Lemma 6, we obtain

$$\mathrm{rank}(T) = \mathrm{rank}(L_{[T]_\beta^\gamma}) = \mathrm{rank}([T]_\beta^\gamma)$$

as desired. $\square$

- **Example** Let $T : P_3(\mathbf{R}) \to P_3(\mathbf{R})$ be the linear transformation

$$Tf := f - xf',$$

thus for instance

$$Tx^2 = x^2 - x(2x) = -x^2.$$

To find the rank of this operator, we let $\beta := (1, x, x^2, x^3)$ be the standard basis for $P_3(\mathbf{R})$. A simple calculation shows that

$$T1 = 1; \quad Tx = 0; \quad Tx^2 = -x^2; \quad Tx^3 = -2x^3,$$

so

$$[T]_\beta^\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

This matrix clearly has rank 3 (row operations can convert it to row-echelon form with three non-zero rows), so T has rank 3.

- The rank of a matrix measures, in some sense, how close to zero (or how "degenerate") a matrix is; the only matrix with rank 0 is the 0 matrix (why). The largest rank an $m \times n$ matrix can have is $\min(m, n)$ (why? See Lemma 2 and Corollary 8). For instance, a 3×5 matrix can have rank at most 3.

* * * * *

Inverting matrices via row operations

- Proposition 5 has the following consequence.

- **Lemma 10.** Let A be an $n \times n$ matrix. Then A is invertible if and only if it is the product of elementary $n \times n$ matrices.

- **Proof** First suppose that A is the product of elementary matrices. We already know that every elementary matrix is invertible; also from Q6 from the Week 4 homework, we know that the product of two invertible matrices is also invertible. Applying this fact repeatedly we see that A is also invertible.

- Now suppose that A is invertible, thus $L_A : \mathbf{R}^n \to \mathbf{R}^n$ is invertible, and in particular is onto. Thus $R(L_A) = \mathbf{R}^n$, and so $\operatorname{rank}(L_A) = n$, and so A itself must have rank n. By Proposition 5 we thus have

$$A = E_1 E_2 \dots E_a I_n F_1 \dots F_b$$

where $E_1, \dots, E_a, F_1, \dots, F_b$ are elementary $n \times n$ matrices. Since the identity matrix I_n cancels out, we are done. $\square$

- This gives us a way to use row operations to invert a matrix A. Suppose we manage to use a sequence of row operations E_1, $E_2, \dots, E_a$ in turn a matrix A into the identity, thus

$$E_a \dots E_2 E_1 A = I.$$

Then by multiplying both sides on the right by A^{-1} we get

$$E_a \dots E_2 E_1 I = A^{-1}.$$

Thus, if we concatenate A and I together, and apply row operations on the concatenated matrix to turn the A component into I, then the I component will automatically turn to A^{-1}. This is a way of computing the inverse of A.

- **Example**. Suppose we want to invert the matrix

$$A := \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

We combine A and the identity I_2 into a single matrix:

$$(A|I_2) = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right).$$

Then we row reduce to turn the left matrix into the identity. For instance, by subtracting three copies of row 1 from row 2 we obtain

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right)$$

and then by adding row 2 to row 1 we obtain

$$\left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & -2 & -3 & 1 \end{array}\right).$$

Dividing the second row by -2 we obtain

$$\left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{array}\right).$$

This the inverse of A is

$$A^{-1} = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix},$$

since the elementary transformations which convert A to I_2, also convert I_2 to A^{-1}.

* * * * *

Determinants

- We now review a very useful characteristic of matrices - the *determinant* of a matrix. The determinant of a square ($n \times n$) matrix is a number which depends in a complicated way on the entries of that matrix. Despite the complicated definition, it has some very remarkable properties, especially with regard to matrix multiplication, and row and column operations. Unfortunately we will not be able to give the proofs of many of these remarkable properties here; the best way to understand determinants is by means of something called *exterior* *algebra*, which is beyond the scope of this course. Without the tools of exterior algebra (in particular, something called a *wedge product*), proving any of the fundamental properties of determinants becomes very messy. So we will settle just for describing the determinant and stating its basic properties.

- The determinant of a 1×1 matrix is just its entry:

$$\det(a) = a.$$

- The 2×2 determinant is given by the formula

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} := ad - bc.$$

- The $n \times n$ determinant is messier, and is defined in the following strange way. For any row i and column j, define the *minor* $\tilde{A}_{ij}$ of an $n \times n$ matrix A to be the $n-1 \times n-1$ matrix which is A with the i^{th} row and j^{th} column removed. For instance, if

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

then

$$\tilde{A}_{11} = \begin{pmatrix} e & f \\ h & i \end{pmatrix}, \quad \tilde{A}_{12} = \begin{pmatrix} d & f \\ g & i \end{pmatrix}, \quad \tilde{A}_{13} = \begin{pmatrix} d & e \\ g & h \end{pmatrix},$$

etc.

This should not be confused with A_{ij}, which is the entry of A in the i^{th} row and j^{th} column. FOr instance, in the above example $A_{11} = a$ and $A_{12} = b$.

- We can now define the $n \times n$ determine recursively in terms of the $n-1$ determinant by what is called the *cofactor expansion*. To find the determinant of an $n \times n$ matrix A, we pick a row i (any row will do) and set

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

For instance, we have

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

or

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

or

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

It seems that this definition depends on which row you use to perform the cofactor expansion, but the amazing thing is that it doesn't! For instance, in the above example, any three of the computations will lead to the same answer, namely

$$aei - ahf - bdi + bgf + cdh - cge.$$

We would like to explain why it doesn't matter which row (or column; see below) to perform cofactor expansion, but it would require one to develop some new material (on the signature of permutations) which would take us far afield, so we will regretfully skip the derivation.

- The quantities $\det(\tilde{A}^{ij})$ are sometimes known as *cofactors*. As one can imagine, this cofactor formula becomes extremely messy for large matrices (to compute the determinant of an $n \times n$ matrix, the above formula will eventually require us to add or subtract $n!$ terms together!); there are easier ways to compute using row and column operations which we will describe below.

- One can also perform cofactor expansion along a column j instead of a row i:
$$\det(A) := \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det(\tilde{A}_{ij}).$$

This ultimately has to do with a symmetry property $\det(A^t) = \det(A)$ for the determinant, although this symmetry is far from obvious given the definition.

- A special case of cofactor expansion: if A has the form

$$A = \begin{pmatrix} c & 0 \ldots 0 \\ \vdots & B \end{pmatrix},$$

where the $0 \ldots 0$ are a string of $n-1$ zeroes, the $\vdots$ represent a column vector of length $n-1$, and B is an $n-1 \times n-1$ matrix, then $\det(A) = c \det(B)$. In particular, from this and induction we see that the identity matrix always has determinant 1: $\det(I_n) = 1$. Also, we see that the determinant of a lower-diagonal matrix is just the product of the diagonal entries; for instance

$$\det(\begin{pmatrix} a & 0 & 0 \\ d & e & 0 \\ g & h & i \end{pmatrix}) = aei.$$

Because of the symmetry property we also know that upper-diagonal matrices work the same way:

$$\det(\begin{pmatrix} a & b & c \\ 0 & e & f \\ 0 & 0 & i \end{pmatrix}) = aei.$$

So in particular, diagonal matrices have a determinant which is just multiplication along the diagonal:

$$\det(\begin{pmatrix} a & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & i \end{pmatrix}) = aei.$$

- An $n \times n$ matrix can be thought of as a collection of n row vectors in $\mathbf{R}^n$, or as a collection of n column vectors in $\mathbf{R}^n$. Thus one can talk about the determinant $\det(v_1, \ldots, v_n)$ of n column vectors in $\mathbf{R}^n$, or the determinant of n row vectors in $\mathbf{R}^n$, simply by arranging those n row or column vectors into a matrix. Note that the order in which we arrange these vectors will be somewhat important.

- Example: the determinant of the two vectors $\begin{pmatrix} a \\ c \end{pmatrix}$ and $\begin{pmatrix} b \\ d \end{pmatrix}$ is $ad - bc$.

- The determinant of n vectors has two basic properties. One is that it is linear in each column separately. What we mean by this is that

$$\det(v_1, \ldots, v_{j-1}, v_j + w_j, v_{j+1}, \ldots, v_n) =$$

$$\det(v_1, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n) + \det(v_1, \ldots, v_{j-1}, w_j, v_{j+1}, \ldots, v_n)$$

  and

$$\det(v_1, \ldots, v_{j-1}, cv_j, v_{j+1}, \ldots, v_n) = c \det(v_1, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n).$$

  This linearity can be seen most easily by cofactor expansion in the column j.

- The other basic property it has is anti-symmetry: if one switches two column vectors (not necessarily adjacent), then the determinant changes sign. For instance, when $n = 5$,

$$\det(v_1, v_5, v_3, v_4, v_2) = -\det(v_1, v_2, v_3, v_4, v_5).$$

  This is not completely obvious from the cofactor expansion definition, although the presence of the factor $(-1)^{i+j}$ does suggest that some sort of sign change might occur when one switches rows or columns. We will not prove this anti-symmetry property here.

- (It turns out that the determinant is in fact the *only* expression which obeys these two properties, and which also has the property that the identity matrix has determinant one. But we will not prove that here).

- The same facts hold if we replace columns by rows; i.e. the determinant is linear in each of the rows separately, and if one switches two rows then the determinant changes sign.

- We now write down some properties relating to how determinants behave under elementary row operations.

- **Property 1** If A is an $n \times n$ matrix, and B is the matrix A with two distinct rows i and j interchanged, then $\det(B) = -\det(A)$. (I.e. row operations of the first type flip the sign of the determinant). This is just a restatement of the antisymmetry property for rows.

- Example:

$$\det \begin{pmatrix} c & d \\ a & b \end{pmatrix} = -\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

- Corollary of Property 1: If two rows of a matrix A are the same, then the determinant must be zero.

- **Property 2** If A is an $n \times n$ matrix, and B is the matrix B but with one row i multiplied by a scalar c, then $\det(B) = c\det(A)$. (I.e. row operations of the second type multiply the determinant by whatever scalar was used in the row operation). This is a special case of the linearity property for the i^{th} row.

- Example:

$$\det \begin{pmatrix} ka & kb \\ c & d \end{pmatrix} = k\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

- Corollary of Property 2: if a matrix A has one of its rows equal to zero, then $\det(A)$ is zero (just apply this Property with $c = 0$).

- **Property 3** If A is an $n \times n$ matrix, and B is the matrix B but with c copies of one row i added to another row j, then $\det(B) = \det(A)$. (I.e. row operations of the third type do not affect the determinant). This is a consequence of the linearity property for the j^{th} row, combined with the Corollary to Property 1 (why?).

- Example:

$$\det \begin{pmatrix} a+kc & b+kd \\ c & d \end{pmatrix} = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

- Similar properties hold for elementary column operations (just replace "row" by "column" throughout in the above three properties).

- We can summarize the above three properties in the following lemma (which will soon be superceded by a more general statement):

- **Lemma 11.** If E is an elementary matrix, then $\det(EA) = \det(E)\det(A)$ and $\det(AE) = \det(A)\det(E)$.

- This is because the determinant of a type 1 elementary matrix is easily seen to be -1 (from Property 1 applied to the identity matrix), the determinant of a type 2 elementary matrix (multiplying a row by c) is easily seen to be c (from Property 2 applied to the identity matrix), and the determinant of a type 3 elementary matrix is easily seen to be 1 (from Property 3 applied to the identity matrix). In particular, elementary matrices always have non-zero determinant (recall that in the type 2 case, c must be non-zero).

- We are now ready to state one of the most important properties of a determinant: it measures how invertible a matrix is.

- **Theorem 12.** An $n \times n$ matrix is invertible if and only if its determinant is non-zero.

- **Proof** Suppose A is an invertible $n \times n$ matrix. Then by Lemma 10, it is the product of elementary matrices, times the identity I_n. The identity I_n has a non-zero determinant: $\det(I_n) = 1$. Each elementary matrix has non-zero determinant (see above), so by Lemma 11 if a matrix has non-zero determinant, then after multiplying on the left or right by an elementary matrix it still has non-zero determinant. Applying this repeatedly we see that A must have non-zero determinant.

  Now conversely suppose that A had non-zero determinant. By Lemma 11, we thus see that even after applying elementary row and column operations to A, one must still obtain a matrix with non-zero determinant. In particular, in row-echelon form A must still have non-zero determinant, which means in particular that it cannot contain any rows which are entirely zero. Thus A has full rank n, which means that $L_A : \mathbf{R}^n \to \mathbf{R}^n$ is onto. But then L_A would also be one-to-one by the dimension theorem - see Lemma 2 of Week 3 notes, hence L_A would be invertible and hence A is invertible. $\square$

- Not only does the determinant measure invertibility, it also measures linear independence.

- **Corollary 13.** Let $v_1, \ldots, v_n$ be n column vectors in $\mathbf{R}^n$. Then $v_1, \ldots, v_n$ are linearly dependent if and only if $\det(v_1, \ldots, v_n) = 0$.

- **Proof** Suppose that $\det(v_1, \ldots, v_n) = 0$, so that by Theorem 12 the $n \times n$ matrix $(v_1, \ldots, v_n)$ is not invertible. Then the linear transformation $L_{(v_1, \ldots, v_n)}$ cannot be one-to-one, and so there is a non-zero vector $\begin{pmatrix} a_1 \\ a_2 \\ \ldots \\ a_n \end{pmatrix}$ in the null space, i.e.

$$(v_1, \ldots, v_n) \begin{pmatrix} a_1 \\ a_2 \\ \ldots \\ a_n \end{pmatrix} = 0$$

  or in other words

$$a_1 v_1 + \ldots + a_n v_n = 0,$$

  i.e. $v_1, \ldots, v_n$ are not linearly independent. The converse statement follows by reversing all the above steps and is left to the reader. □

- Note that if one writes down a typical $n \times n$ matrix, then the determinant will in general just be some random number and will usually not be zero. So "most" matrices are invertible, and "most" collections of n vectors in $\mathbf{R}^n$ are linearly independent (and hence form a basis for $\mathbf{R}^n$, since $\mathbf{R}^n$ is n-dimensional).

- Properties 1,2,3 also give a way to compute the determinant of a matrix - use row and column operations to convert it into some sort of triangular or diagonal form, for which the determinant is easy to compute, and then work backwards to recover the original determinant of the matrix.

- **Example**. Suppose we wish to compute the determinant of

$$A := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 4 \\ 3 & 2 & 1 \end{pmatrix}.$$

We perform row operations. Subtracting two copies of row 1 from row 2 and using Property 3, we obtain

$$\det(A) = \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & -2 \\ 3 & 2 & 1 \end{pmatrix}.$$

Similarly subtracting three copies of row 1 from row 2, we obtain

$$\det(A) = \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & -2 \\ 0 & -4 & -8 \end{pmatrix}.$$

Dividing the third row by $-1/4$ using Property 2, we obtain

$$\frac{-1}{4} \det(A) = \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & -2 \\ 0 & 1 & 2 \end{pmatrix}$$

which after swapping two rows using Property 1, becomes

$$\frac{1}{4} \det(A) = \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & -2 \end{pmatrix}.$$

But the right-hand side is triangular and has a determinant of -2. Hence $\frac{1}{4}\det(A) = -2$, so that $\det(A) = -8$. (One can check this using the original formula for determinant. Which approach is less work? Which approach is less prone to arithmetical error?)

- We now give another important property of a determinant, namely its multiplicative properties.

- **Theorem 14.** If A and B are $n \times n$ matrices, then $\det(AB) = \det(A)\det(B)$.

- **Proof** First suppose that A is not invertible. Then L_A is not onto (cf. Lemma 2 of Week 3 notes), which implies that $L_A L_B$ is not onto (why? Note that the range of $L_A L_B$ must be contained in the range of L_A), so that AB is not invertible. Then by Theorem 12, both sides of

$\det(AB) = \det(A)\det(B)$ are zero, and we are done. Similarly, suppose that B is not invertible. Then L_B is not one-to-one, and so $L_A L_B$ is not one-to-one (why? Note that the null space of L_B must be contained in the null space of $L_A L_B$). So AB is not invertible. Thus both sides are again zero.

- The only remaining case is when A and B are both invertible. By Lemma 10 we may thus write

$$A = E_1 E_2 \ldots E_a; \quad B = F_1 F_2 \ldots F_b$$

where $E_1, \ldots, E_a, F_1, \ldots, F_b$ are elementary matrices. By many applications of Lemma 11 we thus have

$$\det(A) = \det(E_1)\det(E_2)\ldots\det(E_a)$$

and

$$\det(B) = \det(F_1)\det(F_2)\ldots\det(F_b).$$

But also

$$AB = E_1 \ldots E_A F_1 \ldots F_b$$

and so by taking det of both sides and using Lemma 11 many times again we obtain

$$\det(AB) = \det(E_1)\ldots\det(E_A)\det(F_1)\ldots\det(F_b)$$

and by combining all these equations we obtain $\det(AB) = \det(A)\det(B)$ as desired. □

- **Warning:** The corresponding statement for addition is not true in general, i.e. $\det(A+B) \neq \det(A) + \det(B)$ in general. (Can you think up a counterexample? Even for diagonal matrices one can see this will not work. On the other hand, we still have linearity in each row and column).

- Note that Theorem 14 supercedes Lemma 11, although we needed Lemma 11 as an intermediate step to prove Theorems 12 and 14.

- (Optional) Remember the symmetry property $\det(A^t) = \det(A)$ we stated earlier? It can now be proved using the above machinery. We sketch a proof as follows. First of all, if A is non-invertible, then A^t is also non-invertible (why?), and so both sides are zero. Now if A is invertible, then by Lemma 10 it is the product of elementary matrices:

  $$A = E_1 E_2 \ldots E_a$$

  and so

  $$\det(A) = \det(E_1) \ldots \det(E_a).$$

  On the other hand, taking transposes (and recalling that transpose reverses multiplication order) we obtain

  $$A^t = E_a^t \ldots E_2^t E_1^t$$

  and so

  $$\det(A^t) = \det(E_a^t) \ldots \det(E_1^t).$$

  But a direct computation (checking the three types of elementary matrix separately) shows that $\det(E^t) = \det(E)$ for every elementary matrix, so

  $$\det(A^t) = \det(E_a) \ldots \det(E_1).$$

  Thus $\det(A^t) = \det(A)$ as desired.

* * * * *

Geometric interpretation of determinants (optional)

- This material is optional, and is also not covered in full detail. It is intended only for those of you who are interested in the geometric provenance of determinants.

- Up until now we've treated the determinant as a mysterious algebraic expression which has a lot of remarkable properties. But we haven't delved much into what the determinant actually means, and why we have any right to have such a remarkable characteristic of matrices. It turns out that the determinant measures something very fundamental to the geometry of $\mathbf{R}^n$, namely n-dimensional volume. The one caveat is that determinants can be either positive or negative, while volume can only be positive, so determinants are in fact measuring *signed volume* - volume with a sign. (This is similar to how a definite integral $\int_a^b f(x)\,dx$ can be negative if f dips below the x axis, even though the "area under the curve" interpretation of $f(x)$ seems to suggest that integrals must always be positive).

- Let's begin with $\mathbf{R}^1$. The determinant $\det(v_1)$ of a single vector $v_1 = (a)$ in $\mathbf{R}^1$ is of course a, which is plus or minus the length $|a|$ of that vector; plus if the vector is pointing right, and minus if the vector is pointing left. In the degenerate case $v_1 = 0$, the determinant is of course zero.

- Now let's look at $\mathbf{R}^2$, and think about the determinant $\det(v_1, v_2)$ of two vectors v_1, v_2 in $\mathbf{R}^2$. This turns out to be (plus or minus) the area of the parallelogram with sides v_1 and v_2; plus if v_2 is anticlockwise of v_1, and minus if v_2 is clockwise of v_1. For instance, $\det((1,0),(0,1))$ is the area of the square with sides $(1,0),(0,1)$, i.e. 1. On the other hand, $\det((0,1),(1,0))$ is -1 because $(0,1)$ is clockwise of $(1,0)$. Similarly, $\det((3,0),(0,1))$ is 3, because the rectangle with sides $(3,0)$, $(0,1)$ has area 3, and $\det((3,1),(0,1))$ is also 3, because the parallelogram with sides $(3,1),(0,1)$ has the same area as the previous rectangle.

- This parallelogram property can be proven using cross products (recall that the cross product can be used to measure the area of a parallelogram). It is also interesting to interpret Properties 1, 2, 3 using this area interpretation. Property 1 says that if you swap the two vectors v_1 and v_2, then the sign of the determinant changes. Property 2 says that if you dilate one of the vectors by c, then the area of the parallelogram also dilates by c (note that if c is negative, then the determinant changes sign, even though the area is of course always positive, because you flip the clockwiseness of v_1 and v_2). Property 3 says that if you slide v_2 (say) by a constant multiple of v_1, then the area of the parallelogram doesn't change. (This is consistent with the familiar "base $\times$ height" formula for parallelograms - sliding v_2 by v_1 does not affect either the base or the height).

- Note also that if v_1 and v_2 are linearly dependent, then their parallelogram has area 0; this is consistent with Corollary 13.

- Now let's look at $\mathbf{R}^3$, and think about the determinant $\det(v_1, v_2, v_3)$ of three vectors in $\mathbf{R}^3$. Thus turns out to be (plus or minus) the volume of the *parallelopiped* with sides v_1, v_2, v_3 (you may remember this from Math 32A). To determine plus or minus, one uses the right-hand rule: if the thumb is at v_1 and the second finger is at v_2, and the middle finger is at v_3, then we have a plus sign if this can be achieved using the right hand, and a minus sign if it can be achieved using the left-hand. For instance, $\det((1,0,0),(0,1,0),(0,0,1)) = 1$, but $\det((0,1,0),(1,0,0),(0,0,1)) = -1$. It is an instructive exercise to interpret Properties 1,2,3 using this geometric picture, as well as Corollary 13.

- The two-dimensional rule of "determinant is positive if v_2 clockwise of v_1" can be interpreted as a right-hand rule using a two-dimensional hand, while the one-dimensional rule of "determinant is positive if v_1 is on the right" can be interpreted as a right-hand rule using a one-dimensional hand.

- There is a similar link between determinant and n-dimensional volume in higher dimensions $n \geq 3$, but it is of course much more difficult to visualize, and beyond the scope of this course (one needs some grounding in measure theory, anyway, in order to understand what "n-dimensional volume" means. Also, one needs n-dimensional hands.). But in particular, we see that the volume of a parallelopiped with edges $v_1, \ldots, v_n$ is the *absolute value* of the determinant $\det(v_1, \ldots, v_n)$. (Note that it doesn't particularly matter whether we use row or column vectors here since $\det(A^t) = \det(A)$).

- Let $v_1, \ldots, v_n$ be n column vectors in $\mathbf{R}^n$, so that $(v_1, \ldots, v_n)$ is an $n \times n$ matrix, and consider the linear transformation

$$T := L_{(v_1, \ldots, v_n)}.$$

Observe that

$$T(1, 0, \ldots, 0) = v_1$$
$$T(0, 1, \ldots, 0) = v_2$$
$$\ldots$$

$$T(0, 0, \ldots, 1) = v_n$$

(why is this the case?) So if we let Q be the unit cube with edges $(1, 0, \ldots, 0), \ldots, (0, 0, \ldots, 1)$, then T will map Q to the n-dimensional parallelopiped with vectors $v_1, \ldots, v_n$. (If you are having difficulty imagining n-dimensional parallelopipeds, you may just want to think about the $n = 3$ case). Thus $T(Q)$ has volume $|\det(v_1, \ldots, v_n)|$, while Q of course had volume 1. Thus T *expands volume* by a factor of $|\det(v_1, \ldots, v_n)|$. Thus the *magnitude* $|\det(A)|$ of a determinant measures how much the linear transformation L_A expands (n-dimensional) volume.

- **Example** Consider the matrix

$$A := \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}$$

which as we know has the corresponding linear transformation

$$L_A(x_1, x_2) = (5x_1, 3x_2).$$

This dilates the x_1 co-ordinate by 5 and the x_2 co-ordinate by 3, so area (which is 2-dimensional volume) is expanded by 15. This is consistent with $\det(A) = 15$. Note that if we replace 3 with -3, then the determinant becomes -15 but area still expands by a factor of 15 (why?). Also, if we replace 3 instead with 0, then the determinant becomes 0. What happens to the area in this case?

- **Example** Now consider the matrix

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

which as we know has the corresponding linear transformation

$$L_A(x_1, x_2) = (x_1 + x_2, x_2).$$

This matrix shears the x_1 co-ordinate horizontally by an amount depending on the x_2 co-ordinate, but area is unchanged (why? It has to do with the base $\times$ height formula for area). This is consistent with the determinant being 1.

- So the magnitude of the determinant measures the volume-expanding properties of a linear transformation. The sign of a determinant will measure the *orientation-preserving* properties of a transformation: will a "right-handed" object remain right-handed when one applies the transformation? If so, the determinant is positive; if however right-handed objects become left-handed, then the determinant is negative.

- **Example** The reflection matrix

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

  corresponds to reflection through the x_1-axis:

$$L_A(x_1, x_2) = (x_1, -x_2).$$

  It is clear that a "right-handed" object (which in two-dimensions, means an arrow pointing anti-clockwise) will reflect to a "left-handed" object (an arrow pointing clockwise). This is why reflections have negative determinant.

- This interpretation of determinant, as measuring both the volume expanding and the orientation preserving properties of a transformation, also allow us to interpret Theorem 14 geometrically. For instance, if $T : \mathbf{R}^n \to \mathbf{R}^n$ expands volume by a factor of 4 and flips the orientation (so $\det[T]_\beta^\beta = -4$, where β is the standard ordered basis), and $S : \mathbf{R}^n \to \mathbf{R}^n$ expands volume by a factor of 3 and also flips the orientation (so $\det[S]_\beta^\beta = -3$), then one can now see why ST should expand volume by 12 and preserve orientation (so $\det[ST]_\beta^\beta = +12$).

- We now close with a little lemma that says that to take the determinant of a matrix, it doesn't matter what basis you use.

- **Lemma 15.** If two matrices are similar, then they have the same determinant.

- **Proof** If A is similar to B, then $B = Q^{-1}AQ$ for some invertible matrix Q. Thus by Theorem 13

$$\det(B) = \det(Q^{-1}) \det(A) \det(Q) = \det(A) \det(Q^{-1}) \det(Q)$$

$$= \det(A)\det(Q^{-1}Q) = \det(A)\det(I_n) = \det(A)$$

as desired. $\square$

- **Corollary 16.** Let $T : \mathbf{R}^n \to \mathbf{R}^n$ be a linear transformation, and let β, β' be two ordered bases for $\mathbf{R}^n$. Then the matrices $[T]_\beta^\beta$ and $[T]_{\beta'}^{\beta'}$ have the same determinant.

- **Proof.** From last week's notes we know that $[T]_\beta^\beta$ and $[T]_{\beta'}^{\beta'}$ are similar, and the result follows from Lemma 15. $\square$