Contents

0.1 Determinants, Formally

- In this supplement, we give a formal development of the determinant. Since we will not use determinants except in the context of doing practical calculations, and the general theory of determinants will not be of theoretical use to us, the material has been relegated to this appendix.

0.1.1 Definition of the Determinant

- Definition: If A is an $n \times n$ matrix, we define $A^{(i,j)}$ to be the $(n-1) \times (n-1)$ matrix obtained by deleting the ith row and jth column from A.
  - Example: For $A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$ we have $A^{(1,1)} = \begin{bmatrix} b_2 & b_3 \\ c_2 & c_3 \end{bmatrix}$ and $A^{(2,3)} = \begin{bmatrix} a_1 & a_2 \\ c_1 & c_2 \end{bmatrix}$.
- Definition: The determinant of a square matrix A, denoted $\det(A)$ or $|A|$, is defined inductively. For a 1×1 matrix $[a]$ it is just the constant a. For an $n \times n$ matrix with $n \geq 2$, define $A^{(1,k)}$ to be the matrix obtained from A by deleting the 1st row and kth column. Then we define $\det(A) = \sum_{k=1}^{n} (-1)^{k+1} a_{1,k} \det(A^{(1,k)})$.
  - Example: The determinant $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ is $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$.
  - Example: The determinant $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$ is $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$.

0.1.2 Determinants and Elementary Row Operations

- Our first goal is to characterize how the determinant behaves under the elementary row operations. As seems reasonable by working out a few examples, the row operations change a determinant in the following manner:
  - Interchanging two rows multiplies the determinant by -1.
  - Multiplying all entries in one row by a constant scales the determinant by the same constant.
  - Adding or subtracting a scalar multiple of one row to another leaves the determinant unchanged.
- In order to establish these results, we will first prove the "linearity" property of the determinant. To state this result conveniently, we introduce some (nonstandard) notation:

- Definition: If A is an $n \times n$ matrix, define $A_{(k)}[\mathbf{v}]$ to be the $n \times n$ matrix obtained by replacing the kth row of A with the vector $\mathbf{v}$.
  - Example: For $A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$ we have $A_{(2)}(1,4,7) = \begin{bmatrix} a_1 & a_2 & a_3 \\ 1 & 4 & 7 \\ c_1 & c_2 & c_3 \end{bmatrix}$.

- Theorem (Linearity of the Determinant): If A is an $n \times n$ matrix, then for any vectors $\mathbf{v}$ and $\mathbf{w}$, any $1 \le k \le n$, and any scalar r, we have $\det(A_{(k)}[\mathbf{v} + r\mathbf{w}]) = \det(A_{(k)}[\mathbf{v}]) + r\det(A_{(k)}[\mathbf{w}])$.
  - In other words, this theorem says that the determinant of an $n \times n$ matrix is a linear function of the kth row, when the other rows are fixed.
  - Proof: We use induction on n. For $n = 1$, the result is trivial.
  - Now suppose $n \ge 2$, and let $\mathbf{v} = \langle b_1, \dots, b_n \rangle$ and $\mathbf{w} = \langle c_1, \dots, c_n \rangle$.
  - If $k = 1$, then we can expand directly along the top row to see that

$$\begin{aligned} \det(A_{(1)}[\mathbf{v} + r\mathbf{w}]) &= \sum_{j=1}^{n} (-1)^{j+1} (b_j + rc_j) \det(A^{(1,j)}) \\ \det(A_{(1)}[\mathbf{v}]) &= \sum_{j=1}^{n} (-1)^{j+1} (b_j) \det(A^{(1,j)}) \\ \det(A_{(1)}[\mathbf{w}]) &= \sum_{j=1}^{n} (-1)^{j+1} (c_j) \det(A^{(1,j)}) \end{aligned}$$

    and so clearly $\det(A_{(1)}[\mathbf{v} + r\mathbf{w}]) = \det(A_{(1)}[\mathbf{v}]) + r\det(A_{(1)}[\mathbf{w}])$ as required.
  - If $k > 1$, then we may apply the induction hypothesis to the $(n-1) \times (n-1)$ matrix $A^{(1,j)}$ for each j, to see that it is a linear function of its $(k-1)$st row.
  - Thus, $\det(A^{(1,j)}_{(k-1)}[\mathbf{v} + r\mathbf{w}]) = \det(A^{(1,j)}_{(k-1)}[\mathbf{v}]) + r\det(A^{(1,j)}_{(k-1)}[\mathbf{w}])$. (The notation means that we have first deleted the 1st row and jth column of A, and then replaced the $(k-1)$st row with the appropriate vector.)
  - Then, by expanding along the top row, we have

$$\begin{aligned} \det(A_{(k)}[\mathbf{v} + r\mathbf{w}]) &= \sum_{j=1}^{n} (-1)^{j+1} a_{1,j} \det(A^{(1,j)}_{(k-1)}[\mathbf{v} + r\mathbf{w}]) \\ &= \sum_{j=1}^{n} (-1)^{j+1} a_{1,j} \left[\det(A^{(1,j)}_{(k-1)}[\mathbf{v}]) + r\det(A^{(1,j)}_{(k-1)}[\mathbf{w}]) \right] \\ &= \left[\sum_{j=1}^{n} (-1)^{j+1} a_{1,j} \det(A^{(1,j)}_{(k-1)}[\mathbf{v}]) \right] + r \left[\sum_{j=1}^{n} (-1)^{j+1} a_{1,j} \det(A^{(1,j)}_{(k-1)}[\mathbf{w}]) \right] \\ &= \det(A_{(k)}[\mathbf{v}]) + r\det(A_{(k)}[\mathbf{w}]) \end{aligned}$$

    as required.

- Corollary (Row of Zeroes): If a matrix has a row of all zeroes, then its determinant is zero.
  - Proof: If the kth row of A is the zero vector, apply the linearity result with $\mathbf{v} = \mathbf{w} = \mathbf{0}$.
  - This yields $\det(A) = \det(A_{(k)}[\mathbf{0}]) = \det(A_{(k)}[\mathbf{0}]) + r\det(A_{(k)}[\mathbf{0}]) = (r+1)\det(A)$ for any scalar r. But the only way this can happen is when $\det(A) = 0$.

- Corollary (Scaling a Row): Scaling a row by c scales the determinant by c.
  - Proof: Apply the linearity result with $\mathbf{v} = \mathbf{0}$ and $\mathbf{w}$ is the kth row of A: we obtain $\det(A_{(k)}[r\mathbf{w}]) = \det(A_{(k)}[\mathbf{0}]) + r\det(A_{(k)}[\mathbf{w}]) = 0 + r\det(A) = r\det(A)$.

- Next, we will prove that interchanging two rows in a matrix will scale its determinant by -1. To do this we will require the following calculation, which is simply the expansion of the determinant along the top two rows:

- Proposition: For an $n \times n$ matrix A with $n \geq 3$, let $A_{[[i,j]]}$ be the matrix obtained by deleting the first two rows, and the ith and jth columns, of A. Then $\det(A) = \sum_{1 \leq i < j \leq n} (-1)^{i+j+1}(a_{1,i}a_{2,j} - a_{2,i}a_{1,j}) \det(A_{[[i,j]]})$. In particular, interchanging the first two rows of A scales $\det(A)$ by -1.

  - Proof: By definition, $\det(A) = \sum_{j=1}^{n} (-1)^{j+1} a_{1,j} \det(A^{(1,j)})$.
  - Furthermore, $\det(A^{(1,j)}) = \sum_{i=1}^{j-1} (-1)^{i+1} a_{2,i} \det(A_{[[i,j]]}) + \sum_{i=j+1}^{n} (-1)^{i+1} a_{2,i} \det(A_{[[i,j]]})$.
  - Plugging in the formulas for $\det(A^{(1,j)})$ into $\det(A)$, we obtain

$$\begin{aligned} \det(A) &= \sum_{j=1}^{n} (-1)^{j+1} a_{1,j} \left[\sum_{i=1}^{j-1} (-1)^{i+1} a_{2,i} \det(A_{[[i,j]]}) + \sum_{i=j+1}^{n} (-1)^{i+1} a_{2,i} \det(A_{[[i,j]]}) \right] \\ &= \sum_{1 \leq i < j \leq n} (-1)^{i+j+1}(a_{1,i}a_{2,j} - a_{2,i}a_{1,j}) \det(A_{[[i,j]]}) \end{aligned}$$

    after rearranging the summations.
  - For the second statement, observe that interchanging the first two rows will reverse the order of the terms $a_{1,i}a_{2,j}$ and $a_{2,i}a_{1,j}$ but leave everything else unchanged.

- Theorem (Interchanging Two Rows): Interchanging two rows of a determinant scales the determinant by -1.

  - Proof: We use induction on n. For $n = 2$, we observe that $\begin{vmatrix} p & q \\ r & s \end{vmatrix} = ps - qr = -\begin{vmatrix} r & s \\ p & q \end{vmatrix}$.
  - Now suppose $n \geq 3$ and that we are interchanging the ath and bth rows of A, where $a < b$, and take B to be the matrix obtained by interchanging the ath and bth rows of A.
  - If $a = 1$ and $b = 2$, then by the result above, interchanging the 1st and 2nd rows of A scales its determinant by -1, and we are done.
  - If $a = 1$ and $b > 2$, first interchange the 1st and 2nd rows of A to obtain the matrix C.
    - Then $\det(C) = \sum_{j=1}^{n} (-1)^{j+1} a_{1,j} \det(C^{(1,j)})$.
    - By the induction hypothesis, interchanging the 1st and $(b-1)$st rows of $C^{(1,j)}$ scales $\det(C^{(1,j)})$ by -1, so $\det(C^{(1,j)}) = -\det(D^{(1,j)})$, where D is the matrix whose first row is the second row of A, whose second row is the bth row of A, and whose bth row is the first row of A.
    - Then $\det(A) = -\det(C) = -\sum_{j=1}^{n} (-1)^{j+1} a_{1,j} \det(C^{(1,j)}) = \sum_{j=1}^{n} (-1)^{j+1} a_{1,j} \det(D^{(1,j)}) = \det(D)$.
    - But now $\det(D) = -\det(E)$ where E is obtained by interchanging the first two rows of D: this is the same matrix obtained by interchanging the first row of A with the bth row, so we conclude $\det(A) = -\det(E)$.
  - If $a, b > 1$, first write $\det(A) = \sum_{j=1}^{n} (-1)^{j+1} a_{1,j} \det(A^{(1,j)})$.
    - By the induction hypothesis, interchanging the $(a-1)$st and $(b-1)$st rows of $A^{(1,j)}$ scales $\det(A^{(1,j)})$ by -1, so $\det(A^{(1,j)}) = -\det(B^{(1,j)})$.
    - Then $\det(A) = \sum_{j=1}^{n} (-1)^{j+1} a_{1,j} \det(A^{(1,j)}) = -\sum_{j=1}^{n} (-1)^{j+1} a_{1,j} \det(B^{(1,j)}) = -\det(B)$.

- We have shown in all cases that interchanging the ath row with the bth row of A scales the determinant by -1, so we are done.

- Using the result above we can now prove that the determinant of a matrix with two equal rows is zero:
- Proposition (Two Equal Rows): A matrix with two equal rows has determinant zero.
  - Proof: We showed earlier that $\det(A) = \sum_{1 \le i < j \le n} (-1)^{i+j+1}(a_{1,i}a_{2,j} - a_{2,i}a_{1,j}) \det(A_{[[i,j]]})$.
  - If rows 1 and 2 are equal then each term $a_{1,i}a_{2,j} - a_{2,i}a_{1,j}$ is zero, so $\det(A) = 0$.
  - Now suppose that the ath and bth rows are equal. Let B be the matrix obtained by interchanging the ath and 1st rows: then $\det(B) = \pm \det(A)$ (the plus sign will occur if $a = 1$, and the minus sign will occur when $a \neq 1$). Then let C be the matrix obtained by interchanging the bth and 2nd rows; in the same way, $\det(C) = \pm \det(B) = \pm \det(A)$.
  - Then A has equal first and second rows, so $\det(A) = 0$. Thus, $\det(C) = 0$.
  - Remark: A similar argument is to interchange the two equal rows: this would scale the determinant by -1, in which case we would conclude $\det(A) = -\det(A)$ so that $2\det(A) = 0$. However, this will only force $\det(A) = 0$ when $2 \neq 0$ in the scalar field (and there do exist fields in which $2 = 0$). The proof we gave above works in every field.

- We can, at last, show that adding or subtracting a scalar multiple of one row to another leaves the determinant unchanged.
- Theorem (Adding a Multiple of A Row to Another): Adding or subtracting a scalar multiple of one row to another leaves a determinant unchanged.
  - Proof: Suppose that we are adding r times the bth row of A to the kth row of A.
  - Apply the linearity result in the kth row of A, where $\mathbf{v}$ is the kth row of a and $\mathbf{w}$ is the bth row of A.
  - This yields $\det(A_{(k)}[\mathbf{v} + r\mathbf{w}]) = \det(A_{(k)}[\mathbf{v}]) + r\det(A_{(k)}[\mathbf{w}])$.
  - But now notice that $A_{(k)}[\mathbf{w}]$ has both its bth and kth rows equal to $\mathbf{w}$, so by the proposition above, its determinant is zero.
  - Hence $\det(A_{(k)}[\mathbf{v} + r\mathbf{w}]) = \det(A_{(k)}[\mathbf{v}])$, as claimed.

0.1.3 Cofactor Expansions

- Corollary (Linear Dependence of Rows): If the rows of a matrix are linearly dependent, its determinant is zero.
  - Proof: If there is a nontrivial linear dependence among the rows, we can apply elementary row and column operations to create a row of all zeroes. The determinant of the resulting matrix is a nonzero scalar times the original determinant, but the new determinant is zero.

- We can also give an expansion formula for the determinant along an arbitrary row:
- Definition: If A is an $n \times n$ matrix, the (j,k) cofactor of A, $C^{(j,k)}$, is defined to be $(-1)^{j+k}\det(A^{(j,k)})$.
  - Example: The $(2,3)$ cofactor of $\begin{bmatrix} 6 & -1 & 3 \\ 0 & 2 & 0 \\ 3 & 0 & 3 \end{bmatrix}$ is $(-1)^{2+3}\begin{vmatrix} 6 & -1 \\ 3 & 0 \end{vmatrix} = (-1)^5 3 = -3$.

- Theorem (Expansion by Minors): If A is an $n \times n$ matrix, then for any fixed k, $\det(A) = \sum_{j=1}^{n} a_{k,j} C^{(j,k)}$.
  - Proof: Starting with the matrix A, successively interchange rows k and 1, rows 1 and 2, rows 2 and 3, ... , and rows $k-2$ and $k-1$.

- The first row of the resulting matrix B will be the kth row of A, and the remaining rows will be rows 1, 2, ... , $k-1$, $k+1$, ... , n of A, in that order.
- By our results, $\det(B) = (-1)^{k-1}\det(A)$. But now by definition, $\det(B) = \sum_{j=1}^{n}(-1)^{j+1}b_{1,j}\det(B^{(1,j)})$, and $B^{(1,j)}$ is the same as the matrix obtained by deleting the kth row and jth column of A, so $B^{(1,j)} = A^{(k,j)}$.
- So we obtain $\det(A) = \det(B) = (-1)^{k-1}\sum_{j=1}^{n}(-1)^{j+1}a_{k,j}\det(A^{(k,j)}) = \sum_{j=1}^{n}a_{k,j}C^{(j,k)}$, as claimed.

* The statement of the expansion-by-minors formula requires some unpacking.
  - Essentially, the idea is that we can compute the determinant by expanding along any row, rather than just the first row (as in the original definition), or along any column. The calculation of the determinant this way is called "expansion by minors".
  - For example, expanding along the second row yields $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = -b_1\begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix} + b_2\begin{vmatrix} a_1 & a_3 \\ c_1 & c_3 \end{vmatrix} - b_3\begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix}$.
  - The only difficulty is remembering which terms have which sign (plus or minus). Each term has a particular sign based on its location in the matrix, as follows: the $(1,1)$ entry has a plus sign, and the remaining elements are filled in in an alternating "checkerboard" pattern: $\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$.

0.1.4 Computing Determinants

* The fundamental properties of the determinant we have derived so far are as follows:
  - The determinant is a linear function of any row, when the other rows are fixed.
  - Interchanging two rows multiplies the determinant by -1.
  - Multiplying all entries in one row by a constant scales the determinant by the same constant.
  - Adding or subtracting a scalar multiple of one row to another leaves the determinant unchanged.
  - We can compute a determinant using "expansion by minors" along any row.
* When choosing a row to expand along, it is best to choose one with many zeroes, as this will reduce the number of smaller determinants that need to be evaluated.
* <u>Example</u>: Find the determinant $\begin{vmatrix} 1 & -1 & 6 & 6 \\ 2 & 4 & 4 & 7 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 8 & 4 \end{vmatrix}$.
  - Using an expansion along the third row, we see that $\begin{vmatrix} 1 & -1 & 6 & 6 \\ 2 & 4 & 4 & 7 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 8 & 4 \end{vmatrix} = 3\begin{vmatrix} 1 & -1 & 6 \\ 2 & 4 & 7 \\ 0 & 0 & 4 \end{vmatrix}$.
  - Expanding along the third row again yields $3\begin{vmatrix} 1 & -1 & 6 \\ 2 & 4 & 7 \\ 0 & 0 & 4 \end{vmatrix} = 12\begin{vmatrix} 1 & -1 \\ 2 & 4 \end{vmatrix} = 12(1\cdot 4 - (-1)\cdot 2) = \boxed{72}$.
* Upper and lower-triangular matrices have many zeroes, and we can use expansion by minors to give a simple formula for their determinants:
* <u>Corollary</u> (Upper and Lower-Triangular Determinants): The determinant of any upper-triangular or lower-triangular matrix is equal to the product of the diagonal entries.

- ◦ Proof: Induct on the size of the matrix. The results are both trivial for a 1×1 matrix.
  - ◦ For an $n \times n$ lower-triangular matrix with $n \geq 2$, expand along the first row: this yields $a_{1,1}$ times the determinant of the lower right $(n-1) \times (n-1)$ matrix, which by the inductive hypothesis is the product of the remaining diagonal entries.
  - ◦ For an $n \times n$ upper-triangular matrix with $n \geq 2$, expand along the bottom row: this yields $(-1)^{2n-2} a_{n,n} = a_{n,n}$ times the determinant of the upper left $(n-1) \times (n-1)$ matrix, which by the inductive hypothesis is the product of the remaining diagonal entries.

- Although expansion by minors (or even just the definition of the determinant) gives a recursive method for computing any $n \times n$ determinant, these methods are quite slow unless the matrix has many zero entries.
  - ◦ Evaluating an $n \times n$ determinant using the definition requires computing n total $(n-1) \times (n-1)$ determinants, each of which requires evaluating $(n-1)$ total $(n-2) \times (n-2)$ determinants.
  - ◦ Continuing in this way, we see that evaluating an $n \times n$ determinant from the definition requires $n!$ total computations (where we say a 1×1 determinant is one computation).
  - ◦ Since $5! = 120$, it is already quite unreasonable to compute a 5×5 determinant by hand using this method, while a 10×10 determinant (note $10! = 3628800$) is entirely out of reach, and even a computer would have trouble with a 30×30 determinant ($30! = 2.65 \cdot 10^{32}$).

- Row-reduction is a far more efficient method for computing large determinants.
  - ◦ It is sufficient to row-reduce a matrix to put it into row-echelon form, since any row-echelon matrix is upper-triangular, and the determinant of an upper triangular matrix is simply the product of the diagonal entries.
  - ◦ Row-reducing an $n \times n$ matrix requires approximately n^2 individual multiplications, although due to the fact that the sizes of the entries in the matrix can grow quite large (if one is trying to avoid introducing denominators), the total number of calculations is on the order of n^3.
  - ◦ For $n = 5$, one can typically row-reduce a matrix by hand to compute a determinant, and even a 10×10 determinant (approximately a few hundred computations) would not be impossible by hand. A computer can easily deal with a 1000×1000 determinant using row-reductions.

- Example: Find the determinant $\begin{vmatrix} 1 & 2 & -1 & 3 \\ 3 & 7 & 0 & 4 \\ -2 & 1 & 1 & 2 \\ -1 & 3 & 16 & 5 \end{vmatrix}$.
  - ◦ By row-reducing, $\begin{vmatrix} 1 & 2 & -1 & 3 \\ 3 & 7 & 0 & 4 \\ -2 & 1 & 1 & 2 \\ -1 & 3 & 16 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 3 & -7 \\ 0 & 5 & -1 & 8 \\ 0 & 5 & 15 & 8 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 3 & -7 \\ 0 & 0 & -16 & 43 \\ 0 & 0 & 0 & 36 \end{vmatrix} = \boxed{-576}$.

0.1.5 Properties and Applications of Determinants

- As another application of the expansion-by-minors formulas, we can essentially give a formula for the inverse of a matrix:

- Corollary (Adjugate Formula): For any matrix A, $A \cdot \mathrm{adj}(A) = \det(A) I_n$, where $\mathrm{adj}(A)$ is the matrix whose (i,j)-entry is given by the (j,i) cofactor $C^{(j,i)}$ of A. In particular, when $\det(A) \neq 0$, $A^{-1} = \dfrac{1}{\det(A)} [\mathrm{adj}(A)]$.
  - ◦ The name $\mathrm{adj}(A)$ is short for adjugate. (Historically, this matrix was sometimes called the "adjoint", but that term is now used to denote a different object.)
  - ◦ Proof: First consider the (k,k) entry in the product $A \cdot \mathrm{adj}(A)$: it is the sum $\displaystyle\sum_{l=1}^{n} a_{k,l} C^{(k,l)}$, which is the cofactor expansion of $\det(A)$ along the kth row. Thus, the (k,k) entry is equal to $\det(A)$, for each k.

- Now consider the (i, j) entry of the product, for $i \neq j$: it is the sum $\sum_{i=1}^{n} a_{i,l} C^{(j,l)}$, which is the expansion of the determinant of the matrix obtained by replacing the jth row of A with the ith one, along its ith row. This determinant is zero since the matrix has two equal rows.
- Hence all of the off-diagonal entries of $A \cdot \text{adj}(A)$ are zero, and the diagonal entries are all equal to $\det(A)$: this means $A \cdot \text{adj}(A) = \det(A) I_n$ as claimed. The second formula then follows immediately.

- Although the adjugate formula does give an explicit formula for the inverse, it is not computationally useful: it is much faster to compute A^{-1} using row reductions.
  - Using the adjugate formula requires finding an $n \times n$ determinant and n^2 total $(n-1) \times (n-1)$ determinants, so even for a 3×3 matrix, the adjugate formula is far less efficient than row reduction.
- Using row reductions, we can prove the other fundamental properties of the determinant.
- Theorem (Inverses and Determinants): A matrix A is invertible precisely when $\det(A) \neq 0$.
  - Proof: Applying an elementary row operation will scale $\det(A)$ by a nonzero constant, and elementary matrices are invertible, so we may apply row operations to A without changing either of the parts of the theorem.
  - So now assume A is in reduced row-echelon form. If A has a row of all zeroes, then A is not invertible and A also has determinant zero, by our earlier results.
  - If A does not have any rows of all zeroes, then A must be the identity matrix, which is invertible and has nonzero determinant.
- The other fundamental property of the determinant is that it is multiplicative.
- Theorem (Multiplication and Determinants): For any $n \times n$ matrices A and B, $\det(AB) = \det(A) \det(B)$.
  - Proof: Observe first that if E is an elementary matrix, then $\det(EB) = \det(E) \det(B)$: this follows from our analysis of the how the elementary row operations affect determinants and the fact that EB is the matrix obtained by applying the elementary row operation to B.
  - If A is invertible, then by our previous analysis we can row-reduce A to the identity matrix, and therefore write $A = E_1 E_2 \cdots E_k$ for elementary matrices $E_1, \ldots, E_k$.
  - Then by repeated application of the result above, $\det(A) = \det(E_1) \det(E_2) \cdots \det(E_k)$.
  - Then, again by the result above, $\det(AB) = \det(E_1 \cdots E_k B) = \det(E_1) \det(E_2) \cdots \det(E_k) \det(B) = \det(A) \det(B)$.
  - Now suppose that A is not invertible. If AB were invertible, then we could write $A^{-1} = B(AB)^{-1}$ and so A would necessarily be invertible (which it is not). Thus, neither A nor AB is invertible, and thus $\det(A) = \det(AB) = 0$, and so $\det(AB) = 0 = \det(A) \det(B)$ once again.
- Corollary (Determinant of Inverse Matrix): If A is invertible, then $\det(A^{-1}) = \dfrac{1}{\det(A)}$.
  - Proof: In $\det(AB) = \det(A) \det(B)$, set $B = A^{-1}$: we get $1 = \det(I_n) = \det(AA^{-1}) = \det(A) \det(A^{-1})$.
- Finally, we show that the determinant of the transpose of a matrix is equal to the determinant of the original matrix.
- Theorem (Determinant of Transpose): For any $n \times n$ matrix A, $\det(A) = \det(A^T)$.
  - Proof: If A is invertible, then we can write $A = E_1 E_2 \cdots E_k$ as a product of elementary (row) matrices.
  - Then $A^T = E_k^T \cdots E_2^T E_1^T$, and so $\det(A^T) = \det(E_k^T) \cdots \det(E_2^T) \det(E_1^T)$.
  - But now it is a simple verification that $\det(E^T) = \det(E)$ for any elementary matrix E. Thus, we conclude $\det(A) = \det(A^T)$.
  - If A is not invertible, then we can write $A = E_1 E_2 \cdots E_k B$ where B is a matrix with a row of all zeroes.

- Then $A^T = B^T E_k^T \cdots E_2^T E_1^T$, so $\det(A^T) = \det(B^T)\det(E_k^T)\cdots\det(E_2^T)\det(E_1^T)$. But B^T is a matrix with a column of all zeroes: such a matrix is not invertible (CB^T would have a column of all zero entries for any matrix C), and therefore has determinant zero.
- Then $\det(A) = 0 = \det(A^T)$ in this case as well.

- The fact that the determinant of a transpose is equal to the determinant of the original matrix means that all of the results about determinants that we proved for rows also hold for columns: interchanging two columns of a matrix scales the determinant by -1, a matrix with two equal columns has determinant zero, we can expand by minors along any column, and so forth.

- As a final application, we can use determinants to solve systems of linear equations whose coefficient matrix is invertible:

- Theorem (Cramer's Rule): If A is an invertible $n \times n$ matrix, then the matrix equation $A\mathbf{x} = \mathbf{c}$ has a unique solution $\mathbf{x} = A^{-1}\mathbf{c}$. Specifically, the ith coordinate of $\mathbf{x}$ is given by $x_i = \dfrac{\det(C_i)}{\det(A)}$, where C_i is the matrix obtained by replacing the ith column of A with the column vector $\mathbf{c}$.
  - Although Cramer's rule is useful theoretically, in practice computing all of the determinants takes much longer than simply computing the inverse matrix A^{-1}.
  - Proof: If A is invertible, then we can multiply both sides of the equation $A\mathbf{x} = \mathbf{c}$ on the left by A^{-1} to see that $\mathbf{x} = A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{c}$. Since $A^{-1} = \dfrac{1}{\det(A)}\operatorname{adj}(A)$, we obtain $\mathbf{x} = \dfrac{\operatorname{adj}(A)\mathbf{c}}{\det(A)}$.
  - The ith coordinate of the matrix product in the numerator is $x_i = \displaystyle\sum_{k=1}^{n} c_i(-1)^{k+i}\det(A^{(i,k)})$, which is the expansion by minors along the ith column for the determinant of the matrix C_i.
  - Therefore, $x_i = \dfrac{\det(C_i)}{\det(A)}$ as claimed.

- Example: Solve the system of equations $3x + z = 0$, $x + 2y - 3z = 1$, $2x - 2y - z = 2$ using Cramer's rule.
  - The coefficient matrix is $C = \begin{bmatrix} 3 & 0 & 1 \\ 1 & 2 & -3 \\ 2 & -2 & -1 \end{bmatrix}$ whose determinant is $\det(C) = -30$.
  - Since this matrix is invertible the system will have a unique solution.
  - We have $C_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 2 & -3 \\ 2 & -2 & -1 \end{bmatrix}$, $C_2 = \begin{bmatrix} 3 & 0 & 1 \\ 1 & 1 & -3 \\ 2 & 2 & -1 \end{bmatrix}$, and $C_3 = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 2 & 1 \\ 2 & -2 & 2 \end{bmatrix}$, and the respective determinants are $\det(C_1) = -6$, $\det(C_2) = 15$, and $\det(C_3) = 18$.
  - Thus, by Cramer's rule, the solution is $(x, y, z) = \left(\dfrac{-6}{-30}, \dfrac{15}{-30}, \dfrac{18}{-30}\right) = \boxed{\left(\dfrac{1}{5}, -\dfrac{1}{2}, -\dfrac{3}{5}\right)}$.

- As with the other formulas involving determinants, Cramer's rule is not particularly useful for practical computation. In addition to requiring the coefficient matrix to be square and invertible, the total amount of computation is much larger.
  - Solving an $n \times n$ system with Cramer's rule requires computing $n + 1$ total $n \times n$ determinants.
  - In comparison, solving the system via row-reduction directly requires only row-reducing one $n \times n$ matrix.

Well, you're at the end of my handout. Hope it was helpful.