Outline

1. Student Learning Objectives
   - SLOs: Properties of the Determinant

2. Properties of the Determinant
   - Transpose, and Linearity...
   - Computational Feasibility of the Determinant

3. More Properties
   - Invertibility, Gauss-Jordan Elimination, Products, Powers
   - Similar Matrices, Inverses

4. Minors, and Laplace (co-factor) Expansion
   - Matrix Minors
   - Laplace (co-factor) Expansion

5. Suggested Problems
   - Suggested Problems 6.2
   - Lecture–Book Roadmap

6. Supplemental Material
   - Metacognitive Reflection
   - Problem Statements 6.2
After this lecture you should:

- Know the Impact of Row Divisions/Swaps/Additions on the value of the Determinant
- Be familiar with computation of the Determinant of Products, Powers, Transposes, and Inverses of matrices: \( \det(AB), \det(A^k), \det(A^T), \text{ and } \det(A^{-1}) \)
- Be able to compute the determinant using
  - Combinatorial “Pattern” approach, \([\text{NOTES 6.1}]\)
  - Row Reductions,
  - Laplace (co-factor) Expansion Method \([\text{“TRADITIONAL” WAY}].\)
Consider the patterns, $P$ which we use to define the determinant; e.g.

$$A = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}, \quad A^T = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

- Clearly, the product of the pattern $\text{prod}(P)$ is preserved in the transposed pattern;
- The number of inversions is preserved: the down-left number and up-right number just switch roles — therefore $\text{sgn}(P^T) = \text{sgn}(P)$.
- Since this is true for all patterns we must have $\det(A^T) = \det(A)$. 
Theorem (Determinant of the Transpose)

If $A$ is a square matrix, then

$$\det(A^T) = \det(A).$$

This means that any property expressed in terms of columns/rows is also true for rows/columns;

e.g. last time we saw that swapping two columns in a 3-by-3 matrix changed the sign of the determinant;

so, by the above, it directly follows that swapping two rows in a 3-by-3 matrix also changes the sign.
Theorem (Linearity of the Determinant in the Columns)

Consider fixed column vectors

$$\vec{v}_1, \ldots, \vec{v}_{i-1}, \vec{v}_{i+1}, \ldots, \vec{v}_n \in \mathbb{R}^n.$$ 

Then the function $T : \mathbb{R}^n \to \mathbb{R}$ defined by

$$T(\vec{x}) = \det \begin{bmatrix} \vec{v}_1 & \ldots & \vec{v}_{i-1} & \vec{x} & \vec{v}_{i+1} & \ldots & \vec{v}_n \end{bmatrix}$$

is a linear transformation.

We can convince ourselves that the theorem is indeed true...
Linearity of the Determinant in the Columns

“Proof”

We note that $\prod(P)$ is linear in all the rows and columns, since it contains exactly one factor from each row/column.

The determinant is a sum of all the $\prod(P_k)$ (with appropriate sign given by $\text{sgn}(P_k)$ — i.e. the determinant is a linear combination of pattern products.

We express $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y})$, and $T(k\vec{x}) = kT(\vec{x})$:

$$
\det \begin{bmatrix} \vec{v}_1 & \ldots & \vec{x} + \vec{y} & \ldots & \vec{v}_n \end{bmatrix} = \\
\det \begin{bmatrix} \vec{v}_1 & \ldots & \vec{x} & \ldots & \vec{v}_n \end{bmatrix} + \\
\det \begin{bmatrix} \vec{v}_1 & \ldots & \vec{y} & \ldots & \vec{v}_n \end{bmatrix}
$$

$$
\det \begin{bmatrix} \vec{v}_1 & \ldots & k\vec{x} & \ldots & \vec{v}_n \end{bmatrix} = k\det \begin{bmatrix} \vec{v}_1 & \ldots & \vec{x} & \ldots & \vec{v}_n \end{bmatrix}
$$

Peter Blomgren, ⟨blomgren.peter@gmail.com⟩
Say you are faced with computing the determinant of a $32 \times 32$ matrix (not very large by modern standards). Using the “pattern”-method, such a computation would require $31 \cdot 32! \approx 8.16 \cdot 10^{36}$ multiplications.
Computational Feasibility of the Determinant

Say you are faced with computing the determinant of a $32 \times 32$ matrix (not very large by modern standards). Using the “pattern”-method, such a computation would require $31 \cdot 32! \approx 8.16 \cdot 10^{36}$ multiplications. Now say you have access to an “exaflop computer” — which can perform $10^{18}$ operations / second; then your computation would only take about $10^{18}$ seconds... how long is that?
Say you are faced with computing the determinant of a $32 \times 32$ matrix (not very large by modern standards). Using the “pattern”-method, such a computation would require $31 \cdot 32! \approx 8.16 \cdot 10^{36}$ multiplications.

Now say you have access to an “exaflop computer” — which can perform $10^{18}$ operations / second; then your computation would only take about $10^{18}$ seconds... how long is that?

- 100 seconds/minute $\sim 10^{16}$ minutes
- 100 minutes/hour $\sim 10^{14}$ hours
- 100 hours/day $\sim 10^{12}$ days
- 1000 days/year $\sim 10^9$ years.

so... only about 7% of the age of the universe.

Perfect midterm question, yeah?!?
How Fast is $10^{18}$ operations/s?

As of November 2017, according to https://www.top500.org/, the fastest supercomputer in the world:

<table>
<thead>
<tr>
<th>Site</th>
<th>National Supercomputing Center in Wuxi China</th>
</tr>
</thead>
<tbody>
<tr>
<td>System</td>
<td>Sunway TaihuLight - Sunway MPP, Sunway</td>
</tr>
<tr>
<td></td>
<td>SW26010 260C 1.45GHz, Sunway NRCPC</td>
</tr>
<tr>
<td>Cores</td>
<td>10,649,600</td>
</tr>
<tr>
<td>Rmax</td>
<td>93,014.6 TFlop/s</td>
</tr>
<tr>
<td>Rpeak</td>
<td>125,435.9 TFlop/s (0.125 $\times 10^{18}$ Flop/s)</td>
</tr>
<tr>
<td>Power</td>
<td>15,371 kW</td>
</tr>
</tbody>
</table>

The most powerful US system is #5 on the list.
Clearly “we” have to figure something out...

- We know that the determinant of an (upper) triangular matrix is the product of the diagonal entries.
- We can take a general matrix and apply row-reductions to generate an upper triangular matrix.
- Can we “glue” these ideas together?

The short answer is “yes.” The slightly longer answer requires us to figure out how the row-reduction operations change (or don’t change) the value of the determinant.
Our three fundamental row-operations are:

1. **Row division**: Dividing a row by a non-zero scalar $k$.
2. **Row swap**: Swapping two rows.
3. **Row addition**: Adding (subtracting) a multiple of one row to another.
Row-Reductions and Determinants: The $2 \times 2$ Case

First, consider the $2 \times 2$ case $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, with $\det(A) = ad - bc$:

1. **Row division:** If $B = \begin{bmatrix} a/k & b/k \\ c & d \end{bmatrix}$, then
   \[
   \det(B) = ad/k - bc/k = \det(A)/k.
   \]

$\Leftrightarrow$ **Scaling of the Determinant**
First, consider the $2 \times 2$ case $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, with $\det(A) = ad - bc$:

1. **Row division:** If $B = \begin{bmatrix} a/k & b/k \\ c & d \end{bmatrix}$, then $\det(B) = ad/k - bc/k = \det(A)/k$.

   $\leadsto$ **Scaling of the Determinant**

2. **Row swap:** If $B = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$, then $\det(B) = cb - da = -\det(A)$.

   $\leadsto$ **Sign Change of the Determinant**
Row-Reductions and Determinants: The $2 \times 2$ Case

First, consider the $2 \times 2$ case $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, with $\det(A) = ad - bc$:

1. **Row division:** If $B = \begin{bmatrix} a/k & b/k \\ c & d \end{bmatrix}$, then
   
   $\det(B) = ad/k - bc/k = \det(A)/k$.
   
   $\Rightarrow$ **Scaling of the Determinant**

2. **Row swap:** If $B = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$, then $\det(B) = cb - da = -\det(A)$.
   
   $\Rightarrow$ **Sign Change of the Determinant**

3. **Row addition:** If $B = \begin{bmatrix} a + kc & b + kd \\ c & d \end{bmatrix}$, then
   
   $\det(B) = (a + kc)d - (b + kd)c = (ad - bc) + k(cd - dc) = \det(A)$.
   
   $\Rightarrow$ **No Change to the Determinant**
Row-Reductions and Determinants: The $n \times n$ Case

Next, we think about extending to the $n \times n$ case:

1. **Row division**: This follows from the linearity!

2a. **Row swap**: This is a bit trickier: consider swapping two adjacent rows, and think about how this impacts ALL the patterns:
   - Each pattern has exactly one member in each row.
   - The left-most entry is either in the top row (no inversion in this piece of the pattern); or in the bottom row (one inversion).

   | Swapping | (no inversion) $\rightarrow$ (one inversion) | (one inversion) $\rightarrow$ (no inversion) |

   - Bottom line: adjacent swaps lead to a sign-flip of the determinant.
Row-Reductions and Determinants

2b. **Row swap**: OK, what about non-adjacent swaps?

- For two rows \( n \) steps apart:
  - The Bottom row can “climb” to the top using \( n \) adjacent row-swaps;
  - then, the former Top row (now in the second row), can “sink” to the bottom location using \( (n - 1) \) additional adjacent row-swaps:

  \[
  \begin{bmatrix}
  \vec{r}_1^T \\
  \vec{r}_2^T \\
  \vdots \\
  \vec{r}_n^T \\
  \vec{r}_{n+1}^T
  \end{bmatrix}
  \bowtie
  \begin{bmatrix}
  \vec{r}_1^T \\
  \vec{r}_2^T \\
  \vdots \\
  \vec{r}_n^T \\
  \vec{r}_{n+1}^T
  \end{bmatrix}
  \bowtie
  \begin{bmatrix}
  \vec{r}_{n+1}^T \\
  \vec{r}_1^T \\
  \vec{r}_2^T \\
  \vdots \\
  \vec{r}_n^T
  \end{bmatrix}
  \]

  - We get \((2n - 1)\) – an odd number – of row-swaps; *i.e.* and odd number of sign-flips; so the value of the determinant flips.

- Bottom line: **Any** row-swap leads to a sign-flip for the determinant.
- This means that if you have two equal rows, and swap them \( \det(A) = -\det(A) \), which means \( \det(A) = 0 \).
Row-Reductions and Determinants

(Scaled) Row addition:

\[
A = \begin{bmatrix}
\vdots \\
\vec{v}_i^T \\
\vdots \\
\vec{v}_j^T \\
\vdots \\
\vdots
\end{bmatrix} \rightarrow B = \begin{bmatrix}
\vdots \\
\vec{v}_i^T \\
\vdots \\
\vec{v}_j^T + k\vec{v}_i^T \\
\vdots \\
\vdots
\end{bmatrix}
\]

Now, we use linearity:

\[
\det\left(\begin{bmatrix}
\vdots \\
\vec{v}_i^T \\
\vec{v}_j^T + k\vec{v}_i^T \\
\vdots
\end{bmatrix}\right) = \det\left(\begin{bmatrix}
\vdots \\
\vec{v}_i^T \\
\vec{v}_j^T \\
\vdots
\end{bmatrix}\right) + k\det\left(\begin{bmatrix}
\vdots \\
\vec{v}_i^T \\
\vdots
\end{bmatrix}\right)
\]

\[
\text{det}(A) = 0
\]
We summarize:

**Theorem (Elementary Row Operations and Determinants)**

a. *If B is obtained from A by dividing a row of A by a scalar k, then*

\[
\det(B) = \frac{1}{k} \det(A)
\]

b. *If B is obtained from A by a row swap, then*

\[
\det(B) = -\det(A)
\]

c. *If B is obtained from A by adding a multiple of a row of A to another row, then*

\[
\det(B) = \det(A)
\]

Analogous results hold for elementary column operations.
Relating $\det(\text{rref}(A))$ and $\det(A)$

Now, if we *in the process* of computing the reduced-row-echelon-form of a matrix $A$

- count the number of row-swaps: $s$, and
- keep track of scalar divisions $k_1, \ldots, k_r$.

then:

$$\det(\text{rref}(A)) = (-1)^s \frac{1}{k_1 k_2 \ldots k_r} \det(A),$$

or

$$\det(A) = (-1)^s k_1 k_2 \ldots k_r \det(\text{rref}(A))$$

This should save us a second or two... Computing $\text{rref}(A)$ of a $32 \times 32$ matrix will require no more than $\approx \frac{32^3}{3} \approx 11,000$ operations (which is slightly smaller than $8.16 \cdot 10^{36}$)
$n^3$ Growth vs. $n!$ Growth

**Figure:** For small problems ($n \leq 4$), it is faster to use the “pattern” method (usually via the Laplace co-factor expansion), but for large ($n \geq 6$) problems using row-reductions is faster.
If $A$ is invertible, then $\text{rref}(A) = I_n$, so that $\det(\text{rref}(A)) = \det(I_n) = 1$, and

$$\det(A) = (-1)^s k_1 k_2 \ldots k_r \neq 0.$$ 

If $A$ is non-invertible, then the last row of $\text{rref}(A)$ is all zeros, and by linearity $\det(\text{rref}(A)) = 0$; so that $\det(A) = 0$.

**Theorem (Invertibility and Determinant)**

A square matrix $A$ is invertible if and only if $\det(A) \neq 0$.

Yes! we can assign a number $\det(A)$ to any square matrix $A$, such that $A$ is invertible if and only if $\det(A) \neq 0$!
If instead of computing $\det(\text{rref}(A))$, we perform elementary row operations on $A$ to transform it into some matrix $B$, where $\det(B)$ is easy to compute; the same rules apply; if we performed $s$ row swaps, and scaled rows by the factors $k_1, \ldots, k_r$, then

$$
\det(A) = (-1)^s k_1 k_2 \ldots k_r \det(B)
$$

Transforming $A$ into upper triangular form $U$ is a popular choice, since

$$
\det(U) = \prod_{k=1}^{n} u_{kk}.
$$
Determinants of Products and Powers

Theorem (Determinants of Products and Powers)

If $A$ and $B$ are $n \times n$ matrices, and $m \in \mathbb{Z}^+$ is a positive integer, then

a. $\det(AB) = (\det(A))(\det(B))$, and

b. $\det(A^m) = (\det(A))^m$.

Proof in the supplements...
(a.)

(i) First we assume $A$ is invertible: The row-operations required to transform $A$ to $I_n$ applied to the augmented system $[ A \mid A B ]$ gives:

$$\text{rref} \left( \begin{bmatrix} A & A B \end{bmatrix} \right) = \begin{bmatrix} I_n & I_n B \end{bmatrix} = \begin{bmatrix} I_n & B \end{bmatrix}$$

i.e. they are equivalent to multiplying both sides of the augmentation by $A^{-1}$.

Keeping track of the $s$ row-swaps, and row divisions (scalings) $k_1, \ldots, k_r$ required to transform $A$ into its RREF-form, we get

$$\det(A) = (-1)^s k_1 k_2 \ldots k_r, \text{ and }$$

$$\det(AB) = (-1)^s k_1 k_2 \ldots k_r \det(B) = (\det(A)) (\det(B)).$$
(a.)

(ii — When A is not invertible): If A is not invertible, then neither is AB (remember \((AB)^{-1} = B^{-1}A^{-1}\) which makes sense if and only if both A and B are invertible) so

\[
(\det(A)) \ (\det(B)) = 0 \cdot \det(B) = \det(AB).
\]

(b.)

Apply part (a.) \((m - 1)\) times to get:

\[
\det(A^m) = \det(AA^{m-1}) = \det(A) \det(A^{m-1}) = \cdots = (\det(A))^m
\]
Example (Similar Matrices)

Consider two similar matrices $A$, $B$; where $S$ is an invertible matrix so that

$$AS = SB.$$ 

The previous theorems then implies that

$$\det(A) \det(S) = \det(S) \det(B),$$ 

so that $\det(A) = \det(B)$. 

Peter Blomgren, ⟨blomgren.peter@gmail.com⟩
Theorem (Determinants of Similar Matrices)

If a matrix $A$ is similar to $B$, then $\det(A) = \det(B)$.

Theorem (Determinant of an Inverse)

If a matrix $A$ is invertible, then

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

Proof: \{Short: relies on fundamental properties/definitions\}. Since $I_n = A^{-1}A$, $1 = \det(I_n) = \det(A^{-1})\det(A)$. 
Revisiting the $3 \times 3$ case, we recall the result of Sarrus’ formula

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

3 (of the 12) multiplications can be “saved” by writing

$$\det(A) = a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{21}(a_{13}a_{32} - a_{12}a_{33}) + a_{31}(a_{12}a_{23} - a_{13}a_{22})$$

We can recognize this as

$$\det(A) = a_{11} \det \begin{bmatrix} * & a_{22} & a_{23} \\ a_{32} & a_{33} & \end{bmatrix} - a_{21} \det \begin{bmatrix} * & a_{12} & a_{13} \\ a_{32} & a_{33} & \end{bmatrix} + a_{31} \det \begin{bmatrix} * & a_{12} & a_{13} \\ a_{22} & a_{23} & \end{bmatrix}$$

$$= a_{11} \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} & \end{bmatrix} - a_{21} \det \begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} & \end{bmatrix} + a_{31} \det \begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} & \end{bmatrix}$$

We introduce a bit of notation and language...
Matrix Minors

Definition (Minors)

For an \( n \times n \) matrix \( A \), let \( A_{ij} \) be the matrix obtained by omitting the \( i^{\text{th}} \) row, and \( j^{\text{th}} \) column of \( A \). The determinant of this \( (n-1) \times (n-1) \) matrix \( A_{ij} \) is called a minor of \( A \).

With this language, the determinant of the \( 3 \times 3 \) matrix \( A \):

\[
\det(A) = a_{11}\det(A_{11}) - a_{21}\det(A_{21}) + a_{31}\det(A_{31}).
\]

This is know as the Laplace expansion, or co-factor expansion of \( \det(A) \) down the first column.

We generalize this common strategy...
Theorem (Laplace (co-factor) Expansion)

We can compute the determinant of an $n \times n$ matrix $A$ by Laplace expansion down any column, or along any row:

- **Expansion down the $j^{th}$ column:**

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

- **Expansion along the $i^{th}$ row:**

  $$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}).$$
Available on Learning Glass videos:
6.2 — 1, 5, 7, 9, 11, 12, 13, 15
Lecture – Book Roadmap

<table>
<thead>
<tr>
<th>Lecture</th>
<th>Book, [GS5–]</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1*</td>
<td>§5.1, §5.2, §5.3</td>
</tr>
<tr>
<td>6.2</td>
<td>§5.1, §5.2, §5.3</td>
</tr>
<tr>
<td>6.3</td>
<td>§5.1, §5.2, §5.3</td>
</tr>
</tbody>
</table>

* Strang does not talk about the combinatorial (pattern) definition of the determinant.
### Metacognitive Exercise — Thinking About Thinking & Learning

<table>
<thead>
<tr>
<th>I know / learned</th>
<th>Almost there</th>
<th>Huh?!?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Right After Lecture</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>After Thinking / Office Hours / SI-session</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>After Reviewing for Midterm/Final</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Peter Blomgren, ⟨blomgren.peter@gmail.com⟩
(6.2.1) Use Gaussian Elimination (Row Reductions) to find the determinant of the matrix

\[
A = \begin{bmatrix}
1 & 1 & 1 \\
1 & 3 & 3 \\
2 & 2 & 5
\end{bmatrix}
\]

(6.2.5) Use Gaussian Elimination (Row Reductions) to find the determinant of the matrix

\[
A = \begin{bmatrix}
0 & 2 & 3 & 4 \\
0 & 0 & 0 & 4 \\
1 & 2 & 3 & 4 \\
0 & 0 & 3 & 4
\end{bmatrix}
\]
Use Gaussian Elimination (Row Reductions) to find the determinants of the matrices

\((6.2.7)\) \(A = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}\), \(B = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 1 & 3 & 3 & 3 \\ 1 & 1 & 1 & 4 & 4 \\ 1 & 1 & 1 & 1 & 5 \end{bmatrix}\).
Consider a $4 \times 4$ matrix $A$ with rows $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$, and $\vec{v}_4$. If $\det(A) = 8$, what is:

\[
\text{(6.2.11)} \quad \det \left( \begin{bmatrix} \vec{v}_1 \\ \vec{v}_2 \\ -9 \vec{v}_3 \\ \vec{v}_4 \end{bmatrix} \right), \quad \text{(6.2.12)} \quad \det \left( \begin{bmatrix} \vec{v}_4 \\ \vec{v}_2 \\ \vec{v}_3 \\ \vec{v}_1 \end{bmatrix} \right),
\]

\[
\text{(6.2.13)} \quad \det \left( \begin{bmatrix} \vec{v}_2 \\ \vec{v}_3 \\ \vec{v}_1 \\ \vec{v}_4 \end{bmatrix} \right), \quad \text{and} \quad \text{(6.2.15)} \quad \det \left( \begin{bmatrix} \vec{v}_1 \\ \vec{v}_1 + \vec{v}_2 \\ \vec{v}_1 + \vec{v}_2 + \vec{v}_3 \\ \vec{v}_1 + \vec{v}_2 + \vec{v}_3 + \vec{v}_4 \end{bmatrix} \right)
\]