### SLOs 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: $\det(AB)$, $\det(A^k)$, $\det(A^T)$, and $\det(A^{-1})$
- Be able to compute the determinant using
  - Combinatorial “Pattern” approach, [Notes 6.1]
  - Row Reductions,
  - Laplace (co-factor) Expansion Method [“Traditional” Way].

### Determinant of the Transpose

Consider the patterns, $P$ which we use to define the determinant; e.g.

\[
A = \begin{bmatrix}
\ast & \ast & \ast & \ast & \ast \\
\ast & \ast & \ast & \ast & \ast \\
\ast & \ast & \ast & \ast & \ast \\
\ast & \ast & \ast & \ast & \ast \\
\ast & \ast & \ast & \ast & \ast \\
\end{bmatrix}, \quad A^T = \begin{bmatrix}
\ast & \ast & \ast & \ast & \ast \\
\ast & \ast & \ast & \ast & \ast \\
\ast & \ast & \ast & \ast & \ast \\
\ast & \ast & \ast & \ast & \ast \\
\ast & \ast & \ast & \ast & \ast \\
\end{bmatrix}
\]

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

Determinant of the Transpose

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.

Linearity of the Determinant in the Columns (Rows)

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 $T : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by
\[
T(\vec{x}) = \det\left( [\vec{v}_1 \ldots \vec{v}_{i-1} \vec{x} \vec{v}_{i+1} \ldots \vec{v}_n] \right)
\]
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 $T(k\vec{x}) = kT(\vec{x})$:
\[
\det([\vec{v}_1 \ldots \vec{x} + \vec{y} \ldots \vec{v}_n]) = \det([\vec{v}_1 \ldots \vec{x} \ldots \vec{v}_n]) + \det([\vec{v}_1 \ldots \vec{y} \ldots \vec{v}_n])
\]
\[
\det([\vec{v}_1 \ldots k\vec{x} \ldots \vec{v}_n]) = k\det([\vec{v}_1 \ldots \vec{x} \ldots \vec{v}_n])
\]

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?

- 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 2016, 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>

Computational Feasibility of the Determinant

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 operation change (or don’t change) the value of the determinant.

Row-Reductions and Determinants

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
   \[
   \det(B) = \frac{ad}{k} - \frac{bc}{k} = \frac{\det(A)}{k}.
   \]
   $\Rightarrow$ *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!

2. **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).
   - After 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

2.b **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:

   $$
   \begin{bmatrix}
   \vec{r}_1^T \\
   \vec{r}_2^T \\
   \vdots \\
   \vec{r}_n^T
   \end{bmatrix}^n \sim
   \begin{bmatrix}
   \vec{r}_{n+1}^T \\
   \vec{r}_1^T \\
   \vdots \\
   \vec{r}_n^T
   \end{bmatrix}^{n-1} \sim
   \begin{bmatrix}
   \vec{r}_1^T \\
   \vec{r}_2^T \\
   \vdots \\
   \vec{r}_n^T
   \end{bmatrix}
   $$

   - We get $(2n-1) = \text{an odd number}$ 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

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.
Row-Reductions and Determinants

\[ \det(\text{rref}(A)) \text{ and } \det(A) \]

Now, if \( \text{we in the process} \) 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 \( \approx 32^3 \approx 11,000 \) operations (which is slightly smaller than \( 8.16 \cdot 10^3 \)).

Invertibility and Determinant

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 \text{ if and only if } \det(A) \neq 0.

Gauss-Jordan Elimination and the Determinant

If instead of computing \( \det(\text{rref}(A)) \), we perform elementary row operations on \( A \) to transform it into some matrix \( B \), and \( \det(B) \) 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. \( \det(A^m) = (\det(A))^m \).

(a.)  
(i — Assume \( A \) is invertible): It is fairly straight-forward to convince yourself that the row-operations required to transform \( A \) to \( I_n \) applied to the augmented system \( \left[ \begin{array}{c|c} A & AB \end{array} \right] \) gives:

\[
\text{rref} \left( \left[ \begin{array}{c|c} A & AB \end{array} \right] \right) = \left[ \begin{array}{c|c} I_n & B \end{array} \right] = \left[ \begin{array}{c|c} I_n & B \end{array} \right]
\]

continued...

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

\[ \det(A^m) = (\det(A))^m \]

Example (Similar Matrices)

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

\[ AS = SB. \]

The previous theorems then says

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

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

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.  
Since \( I_n = A^{-1} A, 1 = \det(I_n) = \det(A^{-1})\det(A) \).
Minors

Ponder, once again, the $3 \times 3$ case, and recall the result of Sarrus' formula

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

We can save 3 (of the 12) multiplications by writing

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

We can recognize this as

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

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

Minors, and Laplace (co-factor) Expansion

Definition (Minors)

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

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

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

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

We generalize this common strategy...

Suggested Problems 6.2

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.