The formula for the determinant
Here are the determinants of a $2 \times 2$ or $3 \times 3$ matrix.

$$\begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix} = x_{11}x_{22} - x_{12}x_{21}. $$

$$\begin{vmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{vmatrix} = x_{11}x_{22}x_{33} - x_{12}x_{21}x_{33} - x_{11}x_{23}x_{32} + x_{12}x_{23}x_{31} + x_{13}x_{21}x_{32} - x_{13}x_{22}x_{31}. $$

The point of this lecture is to give the corresponding formula for an $n \times n$ matrix of any size. I want to emphasize, though, that knowing the properties of determinant is more useful than knowing this formula.
\[ \begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix} = x_{11}x_{22} - x_{12}x_{21}. \]

\[ \begin{vmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{vmatrix} = x_{11}x_{22}x_{33} - x_{12}x_{21}x_{33} - x_{11}x_{23}x_{32} \\
+ x_{12}x_{23}x_{31} + x_{13}x_{21}x_{32} - x_{13}x_{22}x_{31}. \]

The 2 \times 2 determinant is a sum of 2 terms. The 3 \times 3 determinant is the sum of 6 = 1 \times 2 \times 3 terms. The \( n \times n \) determinant will be the sum of \( n! = 1 \times 2 \times 3 \times \cdots \times n \) terms.
$$\begin{vmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{vmatrix} = x_{11}x_{22}x_{33} - x_{12}x_{21}x_{33} - x_{11}x_{23}x_{32}$$

$$+ x_{12}x_{23}x_{31} + x_{13}x_{21}x_{32} - x_{13}x_{22}x_{31}.$$

Each term will look like $x_{1\sigma(1)}x_{2\sigma(2)}\cdots x_{n\sigma(n)}$, where $\sigma(1)$, $\sigma(2)$, $\ldots$, $\sigma(n)$ is some rearrangement* of the numbers 1, 2, $\ldots$, $n$.

* The proper term is *permutation*. 
\[
\text{det } \begin{bmatrix}
  x_{11} & x_{12} & x_{13} \\
  x_{21} & x_{22} & x_{23} \\
  x_{31} & x_{32} & x_{33}
\end{bmatrix} = x_{11}x_{22}x_{33} - x_{12}x_{21}x_{33} - x_{11}x_{23}x_{32} + x_{12}x_{23}x_{31} + x_{13}x_{21}x_{32} - x_{13}x_{22}x_{31}.
\]

Each term will look like \(x_{1\sigma(1)}x_{2\sigma(2)}\cdots x_{n\sigma(n)}\), where \(\sigma(1), \sigma(2), \ldots, \sigma(n)\) is some permutation of the numbers 1, 2, \ldots, \(n\).
\[
\begin{vmatrix}
  x_{11} & x_{12} & x_{13} \\
  x_{21} & x_{22} & x_{23} \\
  x_{31} & x_{32} & x_{33}
\end{vmatrix}
= x_{11}x_{22}x_{33} - x_{12}x_{21}x_{33} - x_{11}x_{23}x_{32} \\
+ x_{12}x_{23}x_{31} + x_{13}x_{21}x_{32} - x_{13}x_{22}x_{31}.
\]

I have told you to sum up one term for each of the \(n!\) permutations of \(\{1, 2, \ldots, n\}\), and that the thing you sum up is \(x_{1\sigma(1)}x_{2\sigma(2)}\cdots x_{n\sigma(n)}\). What remains is to tell you whether to use a plus or a minus sign.
\[
\begin{vmatrix}
  x_{11} & x_{12} & x_{13} \\
  x_{21} & x_{22} & x_{23} \\
  x_{31} & x_{32} & x_{33}
\end{vmatrix}
= x_{11}x_{22}x_{33} - x_{12}x_{21}x_{33} - x_{11}x_{23}x_{32} \\
+ x_{12}x_{23}x_{31} + x_{13}x_{21}x_{32} - x_{13}x_{22}x_{31}.
\]

I have told you to sum up one term for each of the \(n!\) permutations of \(\{1, 2, \ldots, n\}\), and that the thing you sum up is \(x_{1\sigma(1)}x_{2\sigma(2)} \cdots x_{n\sigma(n)}\). What remains is to tell you whether to use a plus or a minus sign.

The rule is that the coefficient of \(x_{1\sigma(1)}x_{2\sigma(2)} \cdots x_{n\sigma(n)}\) is
\[
(-1)^{\#\{(i,j) : 1 \leq i < j \leq n, \sigma(i) > \sigma(j)\}}.
\]
\((-1)^\#\{(i,j) : 1 \leq i < j \leq n, \; \sigma(i) > \sigma(j)\}\).

\[
\begin{bmatrix}
    x_{11} & x_{12} & x_{13} & x_{14} & x_{15} & x_{16} \\
    x_{21} & x_{22} & x_{23} & x_{24} & \boxed{x_{25}} & x_{26} \\
    x_{31} & x_{32} & x_{33} & x_{34} & x_{35} & \boxed{x_{36}} \\
    x_{41} & x_{42} & \boxed{x_{43}} & x_{44} & x_{45} & x_{46} \\
    x_{51} & x_{52} & x_{53} & \boxed{x_{54}} & x_{55} & x_{56} \\
    x_{61} & \boxed{x_{62}} & x_{63} & x_{64} & x_{65} & x_{66}
\end{bmatrix}
\]

\(x_{11}x_{25}x_{36}x_{43}x_{54}x_{62}\)

\(\sigma(1) = 1 \; \sigma(2) = 5 \; \sigma(3) = 6 \; \sigma(4) = 3 \; \sigma(5) = 4 \; \sigma(6) = 2\)
Rescaling a row

\[
\begin{align*}
\det \begin{bmatrix} x_{11} & x_{12} & x_{13} \\
x_{21} & x_{22} & x_{23} \\
x_{31} & x_{32} & x_{33} \end{bmatrix} &= x_{11} x_{22} (c x_{33}) - x_{12} x_{21} (c x_{33}) - x_{11} x_{23} (c x_{32}) \\
& \quad + x_{12} x_{23} (c x_{31}) + x_{13} x_{21} (c x_{32}) - x_{13} x_{22} (c x_{31}) \\
&= c \left( x_{11} x_{22} x_{33} - x_{12} x_{21} x_{33} - x_{11} x_{23} x_{32} + x_{12} x_{23} x_{31} + x_{13} x_{21} x_{32} - x_{13} x_{22} x_{31} \right) = c \det \begin{bmatrix} x_{11} & x_{12} & x_{13} \\
x_{21} & x_{22} & x_{23} \\
x_{31} & x_{32} & x_{33} \end{bmatrix}
\end{align*}
\]
Adding two vectors in one row

\[
\begin{pmatrix}
  x_{11} & x_{12} & x_{13} \\
  x_{21} & x_{22} & x_{23} \\
  y_1+z_1 & y_2+z_2 & y_3+z_3
\end{pmatrix}
\]

\[
\begin{align*}
\det
\begin{pmatrix}
  x_{11} & x_{12} & x_{13} \\
  x_{21} & x_{22} & x_{23} \\
  y_1+z_1 & y_2+z_2 & y_3+z_3
\end{pmatrix}
&= x_{11}x_{22}(y_3+z_3) - x_{12}x_{21}(y_3+z_3) - x_{11}x_{23}(y_2+z_2) \\
&\quad + x_{12}x_{23}(y_1+z_1) + x_{13}x_{21}(y_2+z_2) - x_{13}x_{22}(y_1+z_1) \\
&= \left( x_{11}x_{22}y_3 - x_{12}x_{21}y_3 - x_{11}x_{23}y_2 \\
&\quad + x_{12}x_{23}y_1 + x_{13}x_{21}y_2 - x_{13}x_{22}y_1 \right) \\
&\quad + \left( x_{11}x_{22}z_3 - x_{12}x_{21}z_3 - x_{11}x_{23}z_2 \\
&\quad + x_{12}x_{23}z_1 + x_{13}x_{21}z_2 - x_{13}x_{22}z_1 \right) \\
&= \det \begin{pmatrix}
  x_{11} & x_{12} & x_{13} \\
  x_{21} & x_{22} & x_{23} \\
  y_1 & y_2 & y_3
\end{pmatrix} + \det \begin{pmatrix}
  x_{11} & x_{12} & x_{13} \\
  x_{21} & x_{22} & x_{23} \\
  z_1 & z_2 & z_3
\end{pmatrix}
\]
Switching two rows

\[
\begin{vmatrix}
  x_{11} & x_{12} & x_{13} \\
  x_{21} & x_{22} & x_{23} \\
  x_{31} & x_{32} & x_{33}
\end{vmatrix} = x_{11}x_{22}x_{33} - x_{12}x_{21}x_{33} - x_{11}x_{23}x_{32} \\
+ x_{12}x_{23}x_{31} + x_{13}x_{21}x_{32} - x_{13}x_{22}x_{31}
\]

\[
\begin{vmatrix}
  x_{11} & x_{12} & x_{13} \\
  x_{31} & x_{32} & x_{33} \\
  x_{21} & x_{22} & x_{23}
\end{vmatrix} = x_{11}x_{32}x_{23} - x_{12}x_{31}x_{23} - x_{11}x_{33}x_{22} \\
+ x_{12}x_{33}x_{21} + x_{13}x_{31}x_{22} - x_{13}x_{32}x_{21}
\]

We need to know that switching two rows of a permutation switches its sign.
If we switch rows $a$ and $b$, we always change the contribution from $(a, b)$. In addition, if $a < j < b$ and $\sigma(a) < \sigma(j) < \sigma(b)$ (or $\sigma(a) > \sigma(j) > \sigma(b)$), we change the contribution of $(a, j)$ and $(j, b)$. Thus, we change the contribution of an odd number of pairs.
We have now proved that det:

- Rescales when we rescale a row.
- Adds when we add vectors in one row.
- Switches sign when we switch two rows.
We have now proved that det:

- Rescales when we rescale a row.
- Adds when we add vectors in one row.
- Switches sign when we switch two rows.

Everything else we proved about determinant follows from that. In particular, we now know that

- Determinants can be computed by row reduction.
- $\det A \neq 0$ if and only if $A$ is invertible (if and only if $\text{rref}(A) = \text{Id}$ if and only if the columns of $A$ are linearly independent if and only if . . . )
- $\det(AB) = \det(A) \det(B)$. 
A final warning

Some people learn the following mnemonic for the $3 \times 3$ determinant:

This is dead wrong for $n > 3$
This would give 8 terms. The $4 \times 4$ determinant has $4! = 24$ terms!
We have now proved that det:

- Rescales when we rescale a row.
- Adds when we add vectors in one row.
- Switches sign when we switch two rows.

Everything else we proved about determinant follows from that. In particular, we now know that

- Determinants can be computed by row reduction.
- \( \det A \neq 0 \) if and only if \( A \) is invertible (if and only if \( \text{rref}(A) = \text{Id} \) if and only if the columns of \( A \) are linearly independent if and only if . . . )
- \( \det(AB) = \det(A) \det(B) \).