Matrix Algebra

Addition & Subtraction.

- We can add two matrices iff they are of the same size. The (i, j)th entry of the resulting matrix is the sum of the (i, j)th entries of the two matrices being added.

Example 1.

$$\begin{pmatrix} 3 & -1 & 0 \\ 6 & 2 & -5 \\ 0 & -4 & 8 \end{pmatrix} + \begin{pmatrix} 4 & -1 & 6 \\ 1 & -3 & 9 \\ -2 & 0 & 8 \end{pmatrix} = \begin{pmatrix} 7 & -2 & 6 \\ 7 & -1 & 4 \\ -2 & -4 & 16 \end{pmatrix}$$

♦

- Subtraction is defined similarly.

Scalar Multiplication.

- We can multiply a matrix by an ordinary number, which we call a *scalar*. The product of a matrix A and a scalar r, denoted rA, is the matrix created by multiplying every entry of A by r.

Example 2.

$$2\begin{pmatrix} 3 & -1 & 0 \\ 6 & 2 & -5 \\ 0 & -4 & 8 \end{pmatrix} = \begin{pmatrix} 6 & -2 & 0 \\ 12 & 4 & -10 \\ 0 & -8 & 16 \end{pmatrix}$$

♦

Matrix Multiplication.

- We can also multiply some matrices by other matrices.
- The product AB of two matrices A and B is only defined if the number of columns of A is equal to the number of rows of B.
- To obtain the (i, j)th entry of AB, we multiply the ith row of A and the jth column of B as shown:

$$\begin{aligned} \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{im} \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{mj} \end{pmatrix} &= a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj} \\ &= \sum_{h}^{m} a_{ih}b_{hj}. \end{aligned}$$

- The product of a $k \times m$ and an $m \times n$ matrix is a $k \times n$ matrix.

Example 3.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} r & s & t \\ u & v & w \end{pmatrix} = \begin{pmatrix} ar + bu & as + bv & at + bw \\ cr + du & cs + dv & ct + dw \end{pmatrix}$$

♦

- Note that we *cannot* multiply these matrices in the reverse order. *Order matters for matrix multiplication!*

- The special matrix

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

  has the property that $AI = A$ for any $m \times n$ matrix A and $IB = B$ for any $n \times k$ matrix B. It is called the $n \times n$ *identity matrix*.

Laws of Matrix Algebra.

For any matrices A, B and C, for which the operations are defined, the following properties hold.

- Associativity:

$$\begin{aligned} (A + B) + C &= A + (B + C) \\ (AB)C &= A(BC) \end{aligned}$$

- Commutativity of Addition:

$$A + B = B + A$$

- Distributivity:

$$\begin{aligned} A(B + C) &= AB + AC \\ (A + B)C &= AC + BC \end{aligned}$$

- Matrix multiplication is *not* commutative, i.e. $AB = BA$ does not hold in general, even when both products are defined. For example

$$\begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix},$$

  but

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

The Transpose.

- Another frequently used operation on matrices is the *transpose*. The $n \times m$ transpose of the $m \times n$ matrix A is obtained by interchanging the rows and columns of A and is denoted by A^T or A':

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

- We are 'flipping' the matrix about its diagonal.

- Some properties of the transpose:

$$\begin{aligned} (A+B)^T &= A^T + B^T \\ (A-B)^T &= A^T - B^T \\ (A^T)^T &= A \\ (rA)^T &= rA^T \\ (AB)^T &= B^T A^T \end{aligned}$$

  where A and B are of the same size and r is a scalar. Try proving the last one yourself using the definitions of the transpose and matrix multiplication (no peeking at the textbook!).

- See section 8.2 of S&B for a list of some special matrices.

Systems of Equations in Matrix Form.

- We can now write a system of linear equations in a simplified way using matrices, as

$$Ax = b$$

  where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \; x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \; b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Algebra of Square Matrices.

Definition. The $n \times n$ matrix B is an *inverse* of the $n \times n$ matrix A if $AB = BA = I$. If the matrix B exists, we say that A is *invertible*. ▲

Theorem 1 (S&B 8.5)**.** *An $n \times n$ matrix A can have at most one inverse.*

Proof. Suppose that B and C are both inverses of A. Then

$$C = CI = C(AB) = (CA)B = IB = B. \qquad \blacksquare$$

- We denote this unique inverse matrix by A^{-1}.

Theorem 2 (S&B 8.6)**.** *If an $n \times n$ matrix A is invertible, then it is nonsingular, and the unique solution to the system of linear equations $Ax = b$ is $x = A^{-1}b$.*

Proof. Multiply each side of the system $Ax = b$ by the inverse A^{-1}. Then use laws of matrix multiplication and the definition of the inverse:

$$A^{-1}(Ax) = A^{-1}b \Leftrightarrow (A^{-1}A)x = A^{-1}b \Leftrightarrow Ix = A^{-1}b \Leftrightarrow x = A^{-1}b. \qquad \blacksquare$$

Theorem 3 (S&B 8.7)**.** *If an $n \times n$ matrix A is nonsingular, then it is invertible.*

- Together, these results tell us that a square matrix is invertible iff if is nonsingular.

- The following theorem summarizes some facts about the inverse and the powers of matrices. By definition,

$$A^2 = AA, A^3 = AA^2, \ldots A^{n+1} = AA^n.$$

Theorem 4 (S&B 8.10 and 8.11)**.** *Let A and B be invertible matrices.*

1. $(A^{-1})^{-1} = A$.
2. $(A^T)^{-1} = (A^{-1})^T$.
3. *AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.*
4. *A^m is invertible for any $m \in \mathbb{Z}$ and $(A^m)^{-1} = (A^{-1})^m = A^{-m}$.*
5. *For any $r, s \in \mathbb{Z}$ we have $A^rA^s = A^{r+s}$.*
6. *For any scalar $r \in \mathbb{R}$ we have rA is invertible and $(rA)^{-1} = (1/r)A^{-1}$.*

- One method for finding the inverse (see the proof of S&B theorem 8.7) uses elementary row operations on the augmented matrix $[A|I]$ to reduce it to $[I|A^{-1}]$.

Example 4. Consider any 2×2 matrix:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

- Write down the augmented matrix $[A|I]$:

$$[A|I] = \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right)$$

- If $a = c = 0$, A is singular. So suppose $a \neq 0$. Then we can add $-c/a$ times row 1 to row 2, to obtain the row echelon form:

$$\left(\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & \frac{ad-bc}{a} & -\frac{c}{a} & 1 \end{array} \right)$$

- This shows that A is nonsingular iff $ad - bc \neq 0$.
- Now multiply row 1 by $1/a$ and row 2 by $a/(ad - bc)$ to get

$$\left(\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array} \right).$$

- After adding $-b/a$ times row 2 to row, we obtain the reduced row echelon matrix:

$$\left(\begin{array}{cc|cc} 1 & 0 & \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array} \right).$$

- So the inverse of A is:

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

- Note that if $ad - bc \neq 0$, a and c cannot both be zero. If $a = 0$, we must have $c \neq 0$ and we can go through steps similar to above to obtain an inverse. ♦

The Determinant of a Matrix.

- The *determinant* of a matrix A is denoted $|A|$ or $\det(A)$.
- The determinant is a number we can compute for any square matrix to determine whether a matrix is nonsingular (i.e. the corresponding linear system has exactly one solution for every choice of RHS b).
- A 1×1 matrix (i.e. a scalar) a has an inverse $1/a$ iff $a \neq 0$. So we define the determinant of a as:

$$\det(a) = a$$

- For a 2×2 matrix (see S&B example 8.4):

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} &= a_{11}a_{22} - a_{12}a_{21} \\ &= a_{11}\det(a_{22}) - a_{12}\det(a_{21}). \end{aligned}$$

Example 5.

$$\begin{vmatrix} 5 & -3 \\ 2 & -1 \end{vmatrix} = 5(-1) - (-3)2 = 1$$

♦

Definition. Let A_{ij} be the $(n-1)\times(n-1)$ submatrix formed by deleting row i and column j from the $n\times n$ matrix A. Then the scalar

$$M_{ij} = \det A_{ij}$$

is called the (i,j)th *minor* of A, and the scalar

$$C_{ij} = (-1)^{i+j} M_{ij}$$

is called the (i,j)th *cofactor* of A. ▲

- Remember the signs of the cofactors by arraying them in a matrix:

$$\begin{pmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

- Now we can write the determinant of a 2×2 matrix as

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\det(a_{22}) - a_{12}\det(a_{21}) \;=\; a_{11}C_{11} + a_{12}C_{12}$$
$$= a_{11}M_{11} - a_{12}M_{12}.$$

- The determinant of a 3×3 matrix A is given by:

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \\ &= a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13} \\ &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} \\ &\quad + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \end{aligned}$$

Example 6.

$$\begin{aligned}
\begin{vmatrix} 1 & -2 & 3 \\ -3 & 4 & 2 \\ -1 & -3 & 6 \end{vmatrix} &= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \\
&= 1(-1)^{(1+1)}\begin{vmatrix} 4 & 2 \\ -3 & 6 \end{vmatrix} + (-2)(-1)^{(1+2)}\begin{vmatrix} -3 & 2 \\ -1 & 6 \end{vmatrix} \\
&\quad + 3(-1)^{(1+3)}\begin{vmatrix} -3 & 4 \\ -1 & -3 \end{vmatrix} \\
&= 1(1)[4(6) - (2)(-3)] + (-2)(-1)[-3(6) - 2(-1)] \\
&\quad + (1)3[(-3)(-3) - 4(-1)] \\
&= 1(30) + (-2)(16) + 3(13) \\
&= 37
\end{aligned}$$

♦

Definition. The *determinant* of an $n \times n$ matrix A is given by:

$$\begin{aligned}
\det A &= a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n} \\
&= a_{11}M_{11} - a_{12}M_{12} + \cdots + (-1)^{n+1}a_{1n}M_{1n}.
\end{aligned}$$

▲

- In fact, it turns out that we can expand along any row or column of the matrix to calculate the determinant.
- For example, expanding along the jth column:

$$\begin{aligned}
\det A &= a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} \\
&= (-1)^{j+1}[a_{1j}M_{1j} - a_{2j}M_{2j} + \cdots + (-1)^{n+1}a_{1n}M_{1n}].
\end{aligned}$$

  This can come in useful if it is easier to expand along one row or column because, for instance, it has a lot of zeros.

- Some useful facts can simplify the calculation of the determinant.

Theorem 5 (S&B 9.1,9.2 and 9.5)**.** *Let A and B be square matrices, then*

1. $|A^T| = |A|$.
2. $|AB| = |A||B|$.
3. $|A + B| \neq |A| + |B|$.
4. *If B comes from interchanging two rows or columns of A, then $|B| = -|A|$.*
5. *If A any row (column) is a multiple of another row (column), then $|A| = 0$.*

6. *If B comes from adding a multiple of one row to another row, then $|C| = |A|$.*
7. *If $|A|$ is a lower-triangular, upper-triangular, or diagonal matrix, then its determinant is simply the product of its diagonal entries.*

- Note that we can reduce a matrix to its row echelon form and then use (4)-(7) to calculate the determinant.
- We now put together some facts about square matrices.

Theorem 6 (S&B 8.9 and 9.3)**.** *For any $n \times n$ matrix A, the following statments are equivalent:*

1. *A is invertible.*
2. *A has a right inverse.*
3. *A has a left inverse.*
4. *Every system $Ax = b$ has at least one solution for every b.*
5. *Every system $Ax = b$ has at most one solution for every b.*
6. *A is nonsingular.*
7. *A has maximal rank n.*
8. *A has a nonzero determinant.*

Uses of the Determinant.

Definition. The *adjoint* of a matrix A is the transpose of the matrix of cofactors of A:

$$\begin{pmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{pmatrix}$$

▲

Theorem 7. *Let A be an $n \times n$ nonsingular matrix. Then:*

1. $$A^{-1} = \frac{\text{adj}\, A}{|A|}, \quad \textit{and}$$

2. (*Cramer's rule*) the unique solution $x = (x_1, \cdots, x_n)$ of the system $Ax = b$ is

$$x_i = \frac{|B_i|}{|A|}, \quad \text{for} \quad i = 1, \ldots, n,$$

where B_i is the matrix A with the RHS b replacing the ith column of A.

Example 7. We will calculate the inverse of the matrix A (from our previous example), where

$$A = \begin{pmatrix} 1 & -2 & 3 \\ -3 & 4 & 2 \\ -1 & -3 & 6 \end{pmatrix}.$$

- Since we already calculated $|A|$, all we need to do is calculate $\operatorname{adj} A$
- We calculated $C_{11} = 30$, $C_{12} = 16$ and $C_{13} = 13$ to find $|A|$, we now find the other cofactors.

$$\begin{aligned} C_{21} = & (-1)^{2+1} \begin{vmatrix} -2 & 3 \\ -3 & 6 \end{vmatrix} \\ = & 3 \end{aligned} \qquad \begin{aligned} C_{22} = & (-1)^{2+2} \begin{vmatrix} 1 & 3 \\ -1 & 6 \end{vmatrix} \\ = & 9 \end{aligned}$$

$$\begin{aligned} C_{23} = & (-1)^{2+3} \begin{vmatrix} 1 & -2 \\ -1 & -3 \end{vmatrix} \\ = & 5 \end{aligned}$$

$$\begin{aligned} C_{31} = & (-1)^{3+1} \begin{vmatrix} -2 & 3 \\ 4 & 2 \end{vmatrix} \\ = & -16 \end{aligned} \qquad \begin{aligned} C_{32} = & (-1)^{3+2} \begin{vmatrix} 1 & 3 \\ -3 & 2 \end{vmatrix} \\ = & -11 \end{aligned}$$

$$\begin{aligned} C_{33} = & (-1)^{3+3} \begin{vmatrix} 1 & -2 \\ -3 & 4 \end{vmatrix} \\ = & -2 \end{aligned}$$

- Now arrange these in the matrix of cofactors

$$\begin{pmatrix} 30 & 16 & 13 \\ 3 & 9 & 5 \\ -16 & -11 & -2 \end{pmatrix}$$

- And take the transpose to get the adjoint

$$\operatorname{adj} A = \begin{pmatrix} 30 & 3 & -16 \\ 16 & 9 & -11 \\ 13 & 5 & -2 \end{pmatrix}$$

- Then the inverse is

$$A^{-}1 = \frac{\operatorname{adj} A}{|A|} = \frac{1}{37} \begin{pmatrix} 30 & 3 & -16 \\ 16 & 9 & -11 \\ 13 & 5 & -2 \end{pmatrix}$$

- Finally, check that $AA^{-1} = I$. ♦

Example 8. For 3×3 systems (also, see example 9.4 S&B)

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3, \end{aligned}$$

Cramer's rule states that

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \; x_2 = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \; x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}.$$