Notes on Matrices

Version 3/4/2026.

Matrix Operations

1. (**Rows and columns**) We write R_i or $\mathbf{r}_i$ for the i-th row of a matrix $\mathbf{A}$, and we write C_j or $\mathbf{c}_j$ to mean the j-th column of $\mathbf{A}$. Using the bold letters mean we think of the row/col as a row/col *vector*.
2. (**Identity matrix**) The *identity matrix* $\mathbf{I}_n$ is the $n \times n$ matrix with 1's in the "main diagonal" (top-left to bottom-right) and 0's everywhere else.
$$\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{I}_4 = \dots.$$
We can just write $\mathbf{I}$ if the inteded size is clear from context.
3. (**Zero matrix**) The matrix $\mathbf{0}_n$ is $n \times n$ and filled with all 0's. Like with $\mathbf{I}$, we just write $\mathbf{0}$ if the size is clear from context.
4. (**Scaling**) Done "elementwise":
$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad 3\mathbf{A} = \begin{bmatrix} 3a & 3b \\ 3c & 3d \end{bmatrix}.$$
5. (**Addition**) Done elementwise (only possible if the dimensions match):
$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} + \begin{bmatrix} \alpha & \beta & \gamma \\ \delta & \varepsilon & \phi \end{bmatrix} = \begin{bmatrix} a+\alpha & b+\beta & c+\gamma \\ d+\delta & e+\varepsilon & f+\phi \end{bmatrix}.$$
6. (**Multiplication**, "Column batches/weights") Shown with example. Let
$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ -3 & -1 & 1 \\ 2 & 0 & 1 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 2 & -1 & -2 \\ 0 & 3 & 1 \\ -1 & 2 & 1 \end{bmatrix}$$
Let $\mathbf{c}_1$, $\mathbf{c}_2$, and $\mathbf{c}_3$ be the columns (vectors) of $\mathbf{A}$, so that $\mathbf{A} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 \end{bmatrix}$. Using the *columns of* $\mathbf{B}$ as weights for the *columns of* $\mathbf{A}$:
$$(\mathbf{c}_1 \text{ of } \mathbf{AB}): \qquad \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} = 2\mathbf{c}_1 + 0\mathbf{c}_2 - \mathbf{c}_3 = \begin{bmatrix} 2 \\ -7 \\ 3 \end{bmatrix}$$
$$(\mathbf{c}_2 \text{ of } \mathbf{AB}): \qquad \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix} = -\mathbf{c}_1 + 3\mathbf{c}_2 + 2\mathbf{c}_3 = \begin{bmatrix} 5 \\ 2 \\ 0 \end{bmatrix}$$
$$(\mathbf{c}_3 \text{ of } \mathbf{AB}): \qquad \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = -2\mathbf{c}_1 + \mathbf{c}_2 + \mathbf{c}_3 = \begin{bmatrix} 0 \\ 6 \\ -3 \end{bmatrix}$$

$$\implies \qquad \mathbf{AB} = \begin{bmatrix} 2 & 5 & 0 \\ -7 & 2 & 6 \\ 3 & 0 & -3 \end{bmatrix}.$$

7. (**Multiplication**, "Row batches/weights") Shown by example. Again let

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ -3 & -1 & 1 \\ 2 & 0 & 1 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 2 & -1 & -2 \\ 0 & 3 & 1 \\ -1 & 2 & 1 \end{bmatrix}.$$

Now let $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{r}_3$ be the rows (vectors) of $\mathbf{B}$, so that $\mathbf{B} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{bmatrix}$. Using the *rows of* $\mathbf{A}$ as weights for the *rows of* $\mathbf{B}$:

$$(\mathbf{r}_1 \text{ of } \mathbf{AB}): \qquad \begin{bmatrix} 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{bmatrix} = \mathbf{r}_1 + 2\mathbf{r}_2 + 0\mathbf{r}_3 = \begin{bmatrix} 2 & 5 & 0 \end{bmatrix}$$

$$(\mathbf{r}_2 \text{ of } \mathbf{AB}): \qquad \begin{bmatrix} -3 & -1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{bmatrix} = -3\mathbf{r}_1 - 1\mathbf{r}_2 + \mathbf{r}_3 = \begin{bmatrix} -7 & 2 & 6 \end{bmatrix}$$

$$(\mathbf{r}_3 \text{ of } \mathbf{AB}): \qquad \begin{bmatrix} 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{bmatrix} = 2\mathbf{r}_1 + 0\mathbf{r}_2 + \mathbf{r}_3 = \begin{bmatrix} 3 & 0 & -3 \end{bmatrix}$$

$$\implies \qquad \mathbf{AB} = \begin{bmatrix} 2 & 5 & 0 \\ -7 & 2 & 6 \\ 3 & 0 & -3 \end{bmatrix}.$$

"Algebra" of matrices. Assume all sums and products below are well-defined by using the appropriate sized matrices.

1. (Adding Zero) $\mathbf{A} + \mathbf{0} = \mathbf{A}$.
2. (Commutative Addition) $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$.
3. (Associative Addition) $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$.
4. (Mutliplicative Identity) $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$.
5. (Multiplicative Zero) $\mathbf{A0} = \mathbf{0A} = \mathbf{0}$.
6. (Associative Multiplication) $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$.
7. (Distributive Law) $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$, and $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$.
8. (Scalar Factoring) If k is any scalar, then $\mathbf{A}(k\mathbf{B}) = k\mathbf{AB} = (k\mathbf{A})\mathbf{B}$. Note that this works even if $\mathbf{B}$ is a vector.
9. (Non-commutative Multiplication) $\mathbf{AB} \neq \mathbf{BA}$ usually (rarely it happens).

Inverses

Definition (Invertible vs. non). A *square* matrix $\mathbf{A}$ (say $n \times n$) is invertible (or "*nonsingular*", or "has $\mathbf{A}^{-1}$ defined") if there exists some $n \times n$ matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n$. If this is the case, then $\mathbf{B} = \mathbf{A}^{-1}$. If no such $\mathbf{B}$ exists, then $\mathbf{A}$ is noninvertible (or "*singular*", or "$\mathbf{A}^{-1}$ is undefined").

Some facts about inverses.

1. Only *square* matrices can have inverses.
2. Zero matrices $\mathbf{0}$ are always noninvertible.
3. Identity matrices are invertible, and $\mathbf{I}^{-1} = \mathbf{I}$.
4. If $\mathbf{A}$ is invertible, then $\mathbf{A}^{-1}$ is invertible too, and $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.
5. $\mathbf{A}^{-1}$ is "unique" (if defined): two different matrices cannot both be $\mathbf{A}^{-1}$.
6. If $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$ are defined, and $\mathbf{AB}$ is defined, then $(\mathbf{AB})^{-1}$ is defined, and
$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}, \quad \text{NOT} \quad \mathbf{A}^{-1}\mathbf{B}^{-1}.$$
7. $(\mathbf{A}+\mathbf{B})^{-1} \neq \mathbf{A}^{-1}+\mathbf{B}^{-1}$, just like $\frac{1}{x+y} \neq \frac{1}{x} + \frac{1}{y}$, and there is no formula for $(\mathbf{A}+\mathbf{B})^{-1}$ (other than just computing it directly).
   In fact, $(\mathbf{A}+\mathbf{B})^{-1}$ could be undefined, even if $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$ are both defined: for example, $\mathbf{I}$ is invertible, but $(\mathbf{I}+(-\mathbf{I}))^{-1} = \mathbf{0}^{-1}$ is undefined.

Computing $\mathbf{A}^{-1}$.

If $\mathbf{A}$ is 2×2, we can use a special formula for $\mathbf{A}^{-1}$:

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad \Longrightarrow \qquad \mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\begin{bmatrix} d & -c \\ -b & a \end{bmatrix}.$$

Of course, if $|\mathbf{A}| = 0$ then $\mathbf{A}^{-1}$ is undefined.

For general square $\mathbf{A}$ (this works evens if $\mathbf{A}$ is 2×2), we "attempt to" compute $\mathbf{A}^{-1}$ by

1. Make the augmented matrix $\left[\mathbf{A} \,|\, \mathbf{I}\right]$ (match the size of $\mathbf{I}$ to the size of $\mathbf{A}$).
2. Row reduce until $\mathbf{A}$ is in REF and you have $\left[(\text{REF } \mathbf{A}) \,|\, \mathbf{B}\right]$.

3.a If $(\text{REF } \mathbf{A}) = \mathbf{I}$, then $\mathbf{A}$ is invertible, and $\mathbf{A}^{-1} = \mathbf{B}$.

3.b If $(\text{REF } \mathbf{A}) \neq \mathbf{I}$, then $\mathbf{A}^{-1}$ is undefined.

Solving systems using inverses.

If we are trying to find a solution $\mathbf{x}$ for the vector equation $\mathbf{Ax} = \mathbf{b}$, then, if $\mathbf{A}^{-1}$ is defined, the system/equation has *unique* solution

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}.$$

Matrix transpose

The *transpose* $\mathbf{A}^\intercal$ is made by "swapping" the rows and columns of $\mathbf{A}$. Thus,

$$(\mathbf{r}_1 \text{ of } \mathbf{A}) \text{ is } (\mathbf{c}_1 \text{ of } \mathbf{A}^\intercal), \quad (\mathbf{c}_1 \text{ of } \mathbf{A}) \text{ is } (\mathbf{r}_1 \text{ of } \mathbf{A}^\intercal), \quad \text{and so on.}$$

Properties of transpose.

1. $(\mathbf{A}^\intercal)^\intercal = \mathbf{A}$.
2. $(\mathbf{A}+\mathbf{B})^\intercal = \mathbf{A}^\intercal + \mathbf{B}^\intercal$.
3. $(\mathbf{A}\mathbf{B})^\intercal = \mathbf{B}^\intercal\mathbf{A}^\intercal$, NOT $\mathbf{A}^\intercal\mathbf{B}^\intercal$.

Determinant

Definition (**Determinant,** 2×2)**.**

$$\text{If} \quad \mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \text{then} \quad \det(\mathbf{A}) = |\mathbf{A}| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \overset{\text{DEF}}{=} ad - bc.$$

Definition. (**Cofactor expansion**)

$$\text{Let} \qquad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

1. (**Minor** M_{ij}) The "(i,j)-minor M_{ij} (of $\mathbf{A}$)" is the determinant of the $(n-1)\times(n-1)$ matrix you get from $\mathbf{A}$ by deleting row-i and column-j.
2. (**Cofactor** C_{ij}) Defined $C_{ij} = (-1)^{i+j}M_{ij}$.
3. Pick a row or column, and note the "indices" of the entries of that row/col:

   Ex: if $\mathbf{A}$ is 4×4, then the entries of Row 3 are a_{31}, a_{32}, a_{33}, and a_{34}, so the "indices" for Row 3 are (3,1), (3,2), (3,3) and (3,4).

   Then the cofactor expansion *along row/column #* is then defined to be

   $$a_{(\#1)}C_{(\#1)} + a_{(\#2)}C_{(\#2)} + \cdots + a_{(\#)}C_{(\#)},$$

   where you fill in the indices into the (#) slots.

   The cofactor expansion along Row 3 would be

   $$\begin{aligned} &a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} + a_{34}C_{34} \\ &= +a_{31}M_{31} - a_{32}M_{32} + a_{33}M_{33} - a_{34}M_{34}. \end{aligned}$$

   Note we started with "+" because $(-1)^{3+1} = +1$.

Definition. (**Determinant, general**) For any one (square) matrix $\mathbf{A}$, the cofactor expansions along every row and column of $\mathbf{A}$ all result in the same value. This single value is the determinant of $\mathbf{A}$, denoted by either $\det(\mathbf{A})$ or $|\mathbf{A}|$.

Properties of determinants.

1. $|\mathbf{A}|$ (a.k.a. $\det(\mathbf{A})$) is only defined for square matrices.
2. $|\mathbf{A}|$ can be negative, positive, or 0.

3. $|\mathbf{0}| = 0$, $|\mathbf{I}| = 1$.

4. $\mathbf{A}^{-1}$ is defined *if and only if* $|\mathbf{A}| \neq 0$.

5. If $\mathbf{A}^{-1}$ is defined, then $|\mathbf{A}^{-1}| = \frac{1}{|\mathbf{A}|}$.

6. $|\mathbf{A}^\intercal| = |\mathbf{A}|$.

7. Assuming $\mathbf{A}$ and $\mathbf{B}$ are square and the same size, then $|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|$.

Note: See what is implied by (**7.**):

- *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are invertible, then so is* $\mathbf{AB}$: because both $|\mathbf{A}| \neq 0$ and $|\mathbf{B}| \neq 0$, then

$$|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}| \neq 0,$$

  and $|\mathbf{AB}| \neq 0$ means $\mathbf{AB}$ is invertible.

- *If either of* $\mathbf{A}$ *and* $\mathbf{B}$ *is noninvertible, then* $\mathbf{AB}$ *can't be invertible either*: at least one of $|\mathbf{A}| = 0$ or $|\mathbf{B}| = 0$, so then

$$|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}| = 0,$$

  and $|\mathbf{AB}| = 0$ means $\mathbf{AB}$ is noninvertible.

Definition. A square matrix is (upper/lower) triangular if all the entries (above/below) the "main diagonal" (the top-left-to-bottom-right diagonal) are 0.

Theorem. *The determinant of a triangular matrix (either upper or lower) is the product of the entries on the main diagonal: for example*

$$\det \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} = 1 \cdot 4 \cdot 6 = 24.$$

Changing determinants through row/column ops.

1. (Swapping) The operations $(\mathbf{r}_i \leftrightarrow \mathbf{r}_j)$ and $(\mathbf{c}_i \leftrightarrow \mathbf{c}_j)$ flip the sign of the determinant.

2. (Scaling) Multiplying a *single* row or column by a scalar k, that is, $(\mathbf{r}_i \to k\mathbf{r}_i)$ or $(\mathbf{c}_i \to k\mathbf{c}_i)$, has the effect of multiplying the determinant by k (so $\det \to k \cdot \det$).

Caution: Scaling an *entire* matrix (ex: $\mathbf{A} \to 3\mathbf{A}$) changes *every* row by the scalar, so account for that: for example

$$\det\left((-2) \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} \right) = \begin{vmatrix} -2 & -4 & -6 \\ 0 & -8 & -10 \\ 0 & 0 & -12 \end{vmatrix} = (-2)^3 \begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix} = (-8) \cdot 24 = -192.$$

3. (Adding) The operations $(\mathbf{r}_i \to \mathbf{r}_i + k\mathbf{r}_j)$ and $(\mathbf{c}_i \to \mathbf{c}_i + k\mathbf{c}_j)$ have *no effect* on the determinant.

Definition (Adjoint matrix). If $\mathbf{A}$ is an invertible matrix (so that $|\mathbf{A}| \neq 0$), the adjoint of a matrix $\mathbf{A}$ is the transpose of the matrix that has entries C_{ij}, where C_{ij} is the (i, j)-cofactor of **A**. For example, if **A** is 3×3, then

$$\operatorname{adj}(\mathbf{A}) = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{bmatrix}.$$

If **A** is 4×4, then the adjoint will be 4×4, and so on.

Theorem (**"Inverse Matrix Theorem"**). *If* **A** *is invertible (so that* $|\mathbf{A}| \neq 0$*), then the inverse matrix* $\mathbf{A}^{-1}$ *can be written*

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \cdot \operatorname{adj}(\mathbf{A}).$$

The inverse matrix thm. is useful if you just need to compute a *single entry* of $\mathbf{A}^{-1}$. If you need to compute the full $\mathbf{A}^{-1}$ then use the "$[\mathbf{A}|\,\mathbf{I}\,]$ method".

Cramer's Rule.

Suppose you are given an $n \times n$ matrix **A**, and asked to solve the vector equation $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{x} = (x_1, \ldots, x_n)$. If $|\mathbf{A}| \neq 0$, *Cramer's Rule* gives us a way to compute individual x_i without having to fully solve the system and/or compute $\mathbf{A}^{-1}$.

Note: Because **A** is invertible, we know that there is a *unique* solution to the system, which must be $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

Cramer's Rule:

1. First compute the determinant $|\mathbf{A}|$.

2. Let $\mathbf{a}_1, \ldots, \mathbf{a}_n$ be the *column vectors* of **A**, so that

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}.$$

Then, for $i = 1, \ldots, n$, form the new matrices

$$\mathbf{B}_1 = \begin{bmatrix} \mathbf{b} & \mathbf{a}_2 & \mathbf{a}_3 & \cdots & \mathbf{a}_n \end{bmatrix},$$
$$\mathbf{B}_2 = \begin{bmatrix} \mathbf{a}_1 & \mathbf{b} & \mathbf{a}_3 & \cdots & \mathbf{a}_n \end{bmatrix},$$
$$\cdots\cdots$$
$$\mathbf{B}_n = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \cdots & \mathbf{b} \end{bmatrix}.$$

Thus, $\mathbf{B}_k$ is the matrix where the column $\mathbf{a}_k$ of **A** is replaced with **b**.

3. The individual entries x_k of the unique solution **x** can be computed using

$$x_1 = \frac{|\mathbf{B}_1|}{|\mathbf{A}|}, \quad x_2 = \frac{|\mathbf{B}_2|}{|\mathbf{A}|}, \quad \ldots, \quad x_n = \frac{|\mathbf{B}_n|}{|\mathbf{A}|}.$$