BS1820: Maths and Statistics Foundations for Analytics

Linear Algebra 2

Zhe Liu

Imperial College Business School

Email: zhe.liu@imperial.ac.uk

Outline

2.1 Matrices

Matrices are rectangular arrays of numbers. For example:

- $\mathbf{A} = \begin{bmatrix} 2 & 3 & 7 \\ 1 & 6 & 5 \end{bmatrix}$ is a 2×3 matrix
- $\mathbf{B} = \begin{bmatrix} 2 & 3 & 7 \end{bmatrix}$ is a 1×3 matrix $\equiv$ row vector
- $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$ is an $\mathbf{m} \times \mathbf{n}$ matrix $\in \mathbb{R}^{\mathbf{m} \times \mathbf{n}}$
- $\mathbf{I}_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$ is the $n \times n$ **Identity** matrix

Vectors are clearly also matrices — a special case with one column.

2.2 Matrix Operations: Transpose

Transpose: $\mathbf{A} \in \mathbb{R}^{m \times d}$

$$\mathbf{A}^\top = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1d} \\ a_{21} & a_{22} & \dots & a_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{md} \end{bmatrix}^\top = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1d} & a_{2d} & \dots & a_{md} \end{bmatrix} \in \mathbb{R}^{d \times m}$$

- Transpose of a row vector is a column vector

Examples:

- $\mathbf{A} = \begin{bmatrix} 2 & 3 & 7 \\ 1 & 6 & 5 \end{bmatrix}$ a 2×3 matrix and $\mathbf{A}^\top = \begin{bmatrix} 2 & 1 \\ 3 & 6 \\ 7 & 5 \end{bmatrix}$ a 3×2 matrix
- $\mathbf{v} = \begin{bmatrix} 2 \\ 6 \\ 4 \end{bmatrix}$ a column vector and $\mathbf{v}^\top = \begin{bmatrix} 2 & 6 & 4 \end{bmatrix}$ a row vector

2.2 Matrix Operations: Elementwise

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 7 \\ 1 & 6 & 5 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 8 & 7 \end{bmatrix}$$

Elementwise operations:

1. Addition: $\mathbf{A} + \mathbf{B} = \begin{bmatrix} 3 & 6 & 9 \\ 5 & 14 & 12 \end{bmatrix}$
2. Subtraction: $\mathbf{A} - \mathbf{B} = \begin{bmatrix} 1 & 0 & 5 \\ -3 & -2 & -2 \end{bmatrix}$
3. Scalar multiplication: $\mathbf{5A} = \begin{bmatrix} 10 & 15 & 35 \\ 5 & 30 & 25 \end{bmatrix}$
4. Hadamard product: $\mathbf{A} \circ \mathbf{B} = \begin{bmatrix} 2 & 9 & 14 \\ 4 & 48 & 35 \end{bmatrix}$
5. Hadamard power: $\mathbf{A}^{\circ 2} = \begin{bmatrix} 4 & 9 & 49 \\ 1 & 36 & 25 \end{bmatrix}$

2.2 Matrix Operations: Multiplication

$\mathbf{A} \in \mathbb{R}^{\mathbf{m}\times\mathbf{d}}$, $\mathbf{B} \in \mathbb{R}^{\mathbf{d}\times p}$ then $\mathbf{C} = \mathbf{AB} \in \mathbb{R}^{\mathbf{m}\times p}$ with

$$c_{ij} = \begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{id} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{dj} \end{bmatrix}$$

$$c_{22} = a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} + a_{24}b_{42}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \\ b_{41} & b_{42} & b_{43} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix}$$

2.2 Matrix Operations: Multiplication

Properties:

- Inner product: $\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^\top \mathbf{w}$

  E.g. $\begin{bmatrix} 2 \\ 6 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 7 \\ 2 \\ 8 \end{bmatrix} = \begin{bmatrix} 2 & 6 & 4 \end{bmatrix} \begin{bmatrix} 7 \\ 2 \\ 8 \end{bmatrix} = 2 \times 7 + 6 \times 2 + 4 \times 8 = 58$

- $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$

  E.g. Consider $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. We have

  $$(\mathbf{AB})^T = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}^T = \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} = \mathbf{B}^T \mathbf{A}^T$$

- $\mathbf{AI_n} = \mathbf{I_m A} = \mathbf{A}$ for any $\mathbf{A} \in \mathbb{R}^{\mathbf{m} \times \mathbf{n}}$

2.2 Matrix Operations: Multiplication

Example A production manager is scheduling 2 products using 3 machines. Each product can be manufactured by any of the three machines. The time (in hours) required to produce each product with each machine is given below:

		Machine Hours		
		A	B	C
Product	Product 1	2	3	7
	Product 2	1	6	5

Machine hours

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 7 \\ 1 & 6 & 5 \end{bmatrix}$$

Assume the following production schedule:

		Production (in units)	
		Product 1	Product 2
Machine	A	5	10
	B	10	0
	C	0	3

Production plan

$$\mathbf{B} = \begin{bmatrix} 5 & 10 \\ 10 & 0 \\ 0 & 3 \end{bmatrix}$$

2.2 Matrix Operations: Multiplication

Question: What are the total machine hours spent for each product?

		Machine Hours		
		A	B	C
Product	Product 1	2	3	7
	Product 2	1	6	5

		Production (in units)	
		Product 1	Product 2
Machine	A	5	10
	B	10	0
	C	0	3

		Total Machine Hours Spent
Product	Product 1	2x5+3x10+7x0=40
	Product 2	1x10+6x0+5X3=25

We calculate the total machine hours for each product by multiplying the hours needed by each machine (rows of matrix **A**) and the production made on each machine (columns of matrix **B**): the diagonal elements of **AB**.

2.2 Matrix Operations: Multiplication

Facts:

1. Not any two matrices can be multiplied; the dimensions should be conforming: # columns of the left matrix = # rows of the right matrix.
   E.g. $\begin{bmatrix} 2 & 3 & 7 \\ 1 & 6 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 21 \\ 30 \end{bmatrix}$ but $\begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix} \begin{bmatrix} 2 & 3 & 7 \\ 1 & 6 & 5 \end{bmatrix}$ is infeasible.

2. Matrix multiplication is not commutative, i.e. $\mathbf{AB} \neq \mathbf{BA}$.
   E.g. See above.

3. Matrix multiplication is associative, i.e. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$.

4. Matrix multiplication is distributive wrt addition, e.g. $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$.

Exercises

Question 1: Write the following linear equations in matrix multiplication form.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = y_1 \\ a_{21}x_1 + a_{22}x_2 = y_2 \end{cases}$$

Question 2: Find a matrix that, through matrix multiplication, transforms the following 3×2 matrix to another whose two columns are the sum and difference of the original two columns, respectively.

$$\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \to \begin{bmatrix} 1+4 & 1-4 \\ 2+5 & 2-5 \\ 3+6 & 3-6 \end{bmatrix}$$

Exercises

Question 1: Write the following linear equations in matrix multiplication form.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = y_1 \\ a_{21}x_1 + a_{22}x_2 = y_2 \end{cases}$$

Answer:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

Question 2: Find a matrix that, through matrix multiplication, transforms the following 3×2 matrix to another whose two columns are the sum and difference of the original two columns, respectively.

$$\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \to \begin{bmatrix} 1+4 & 1-4 \\ 2+5 & 2-5 \\ 3+6 & 3-6 \end{bmatrix}$$

Answer: Right multiplied by

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

2.3 Range and Rank

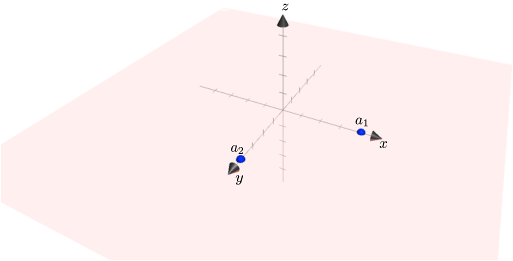
Let $\mathbf{A} \in \mathbb{R}^{m \times n}$, then for any $\mathbf{x} \in \mathbb{R}^n$

$$\mathbf{A}\mathbf{x} = x_1 \mathbf{a}_1 + \cdots + x_n \mathbf{a}_n$$

where $\mathbf{a}_i$ is the i^{th} column of $\mathbf{A}$ and x_i the i^{th} component of $\mathbf{x}$.

Hence $\mathbf{A}\mathbf{x}$ is a linear combination of the columns of $\mathbf{A}$.

Example: Consider $\mathbf{A} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 2}$ with $\mathbf{a}_1 = \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix}$ and $\mathbf{a}_2 = \begin{bmatrix} 0 \\ 4 \\ 0 \end{bmatrix}$.



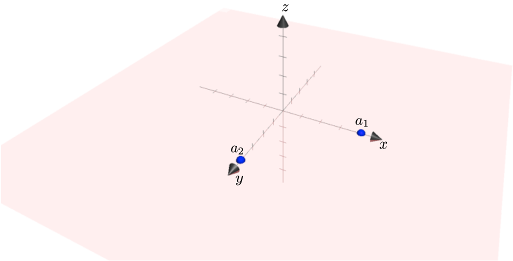
2.3 Range and Rank

Definition. The range of $\mathbf{A} \in \mathbb{R}^{m \times n}$, denoted by $\mathcal{R}(\mathbf{A})$, is the set of vectors in $\mathbb{R}^m$ that can be obtained as a linear combination of the columns of $\mathbf{A}$:

$$\mathcal{R}(\mathbf{A}) := \{\mathbf{A}\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\}.$$

Clearly $\mathcal{R}(\mathbf{A}) = \text{span}(\{\mathbf{a}_1, \ldots, \mathbf{a}_n\})$ and is a subspace of $\mathbb{R}^m$. The dimension of this subspace is the rank of $\mathbf{A}$.

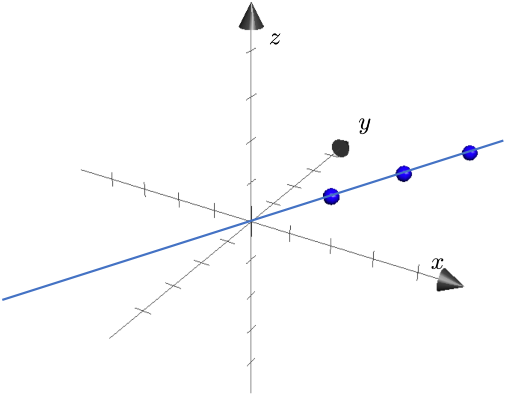
Running example: $\text{span}(\{\mathbf{a}_1, \mathbf{a}_2\}) = \mathbb{R}^2$, $\text{rank}(A) = 2$.



2.3 Range and Rank

Example:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}, \quad \text{range}(\mathbf{A}) = \left\{ \lambda \begin{bmatrix} 1 \\ 2 \end{bmatrix} : \lambda \in \mathbb{R} \right\}, \quad \text{rank}(\mathbf{A}) = 1$$



2.3 Range and Rank

Properties: For $\mathbf{A} \in \mathbb{R}^{m \times n}$,

$$
\begin{aligned}
\textbf{Column rank} \text{ of } \mathbf{A} &= \text{max number of linearly independent columns of } \mathbf{A} \\
&= \text{size of the basis (dimension) of } \textbf{column space } \mathcal{R}(\mathbf{A}) \\
\textbf{Row rank} \text{ of } \mathbf{A} &= \text{max number of linearly independent rows of } \mathbf{A} \\
&= \text{size of the basis (dimension) of } \textbf{row space } \mathcal{R}(\mathbf{A}^T)
\end{aligned}
$$

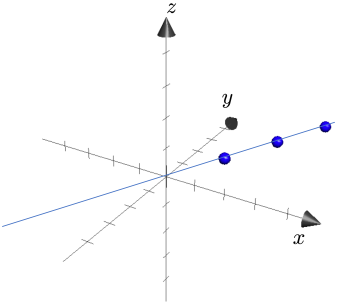
$$
\text{Rank} := \text{Column rank of } \mathbf{A} \equiv \text{Row rank of } \mathbf{A} \leq \min\{m, n\}.
$$

2.3 Range and Rank

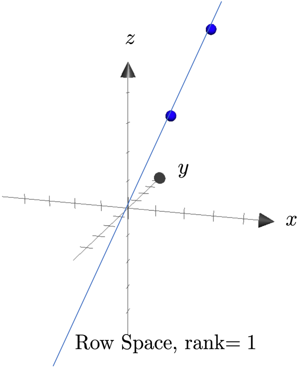
Rank := Column rank of $\mathbf{A} \equiv$ Row rank of $\mathbf{A} \leq \min\{m, n\}$.

Example:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}$$



Column Space, rank= 1



Row Space, rank= 1

2.4 Null Space

Recall: the column vectors of matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ span the column space (range):

$$\mathcal{R}(\mathbf{A}) = \mathsf{span}(\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}) = \{\mathbf{Ax} : \mathbf{x} \in \mathbb{R}^n\}.$$

How about the vectors $\mathbf{x} \in \mathbb{R}^n$ that satisfy $\mathbf{Ax} = \mathbf{0}$?

The space spanned by such vectors is called the **null space (kernel)** of $\mathbf{A}$:

$$\mathcal{N}(\mathbf{A}) := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{Ax} = \mathbf{0}\}$$

Fact: Dimension of column space ($\mathsf{rank}(\mathbf{A})$) + dimension of null space = # of columns:

$$\dim(\mathcal{R}(\mathbf{A})) + \dim(\mathcal{N}(\mathbf{A})) = n.$$

Example:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} \quad \Rightarrow \quad \mathcal{N}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^3 : x_1 + 2x_2 + 3x_3 = 0\}$$

It is a plane! $\dim(\mathcal{N}(\mathbf{A})) = 2$.

2.5 Determinant

Consider a 2×2 square matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

The determinant of $\mathbf{A}$, denoted by $\det(\mathbf{A})$, is defined as

$$\det(\mathbf{A}) := a_{11}a_{22} - a_{12}a_{21}.$$

Any $\mathbf{x} \in \mathbb{R}^2$ is transformed/mapped to $\mathbf{y} \in \mathbb{R}^2$ by $\mathbf{A}\mathbf{x} = \mathbf{y}$. The unit square in $\mathbb{R}^2$ gets mapped to a parallelogram in $\mathbb{R}^2$ **with area** $= |\det(\mathbf{A})|$.

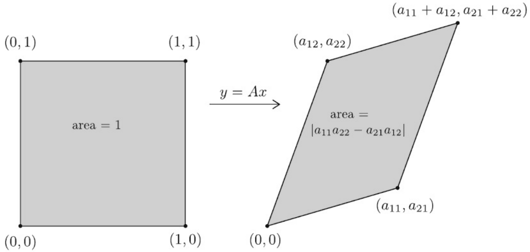


Figure 3.7 from Calafiore and El Ghaoui's's *Optimization Models*: Linear mapping of the unit square.

2.5 Determinant

Definition. For a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, its determinant is defined *inductively* via the formula

$$\det(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\mathbf{A}_{(i,j)}), \quad \text{for any } i = 1, \ldots, n,$$

where $\mathbf{A}_{(i,j)} \in \mathbb{R}^{(n-1)\times(n-1)}$ is $\mathbf{A}$ with its i^{th} row and j^{th} column removed.

- For $\mathbf{A} \in \mathbb{R}^{3\times 3}$, a unit cube $\mapsto$ a parallelotope in $\mathbb{R}^3$ with volume $|\det(\mathbf{A})|$.
- Generally, $\mathbf{A} \in \mathbb{R}^{n \times n}$ maps a unit cube in $\mathbb{R}^n \mapsto$ to a parallelotope in $\mathbb{R}^n$ with volume $|\det(\mathbf{A})|$.

2.5 Determinant: Properties

Definition. A square matrix $\mathbf{A}$ is singular if $\det(\mathbf{A}) = 0$, non-singular otherwise.

For $\mathbf{A}$, $\mathbf{B} \in \mathbb{R}^{n \times n}$ and $\alpha \in \mathbb{R}$ we have:

1. $\det(\mathbf{A}) = \det(\mathbf{A}^\top)$.
   E.g. $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow \det(\mathbf{A}) = a \times d - b \times c$,
   $\mathbf{A}^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \Rightarrow \det(\mathbf{A}^T) = a \times d - b \times c$
2. $\det(\mathbf{AB}) = \det(\mathbf{BA}) = \det(\mathbf{A})\det(\mathbf{B})$.
3. $\det(\alpha\mathbf{A}) = \alpha^n \det(\mathbf{A})$.
4. $\det(\mathbf{I}_n) = 1$.
   E.g. $\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\det(\mathbf{I}_2) = 1 \times 1 - 0 \times 0 = 1$

Exercise: Pick square matrices $\mathbf{A}$ and $\mathbf{B}$ to check properties 2 and 3.

2.6 Matrix Inverse

Definition. We say a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is invertible if there exists a matrix $\mathbf{A}^{-1} \in \mathbb{R}^{n \times n}$ such that

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n.$$

Such a matrix (if exists) is unique and is called the inverse of $\mathbf{A}$.

Theorem. For a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ the following are equivalent:

1. $\mathbf{A}$ is invertible.
2. $\mathbf{A}$ has full rank, i.e. $\text{rank}(\mathbf{A}) = n$.
3. The column space of $\mathbf{A}$ is $\mathbb{R}^n$, i.e. $\mathcal{R}(\mathbf{A}) = \mathbb{R}^n$.
   The null space of $\mathbf{A}$ only contains the zero vector, i.e. $\mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$.
4. The determinant of $\mathbf{A}$ is non-zero (non-singular), i.e. $\det(\mathbf{A}) \neq 0$.