Math Refresher
Winter Institute in Data Science

Ryan T. Moore

2025-12-16

Warming Up

1. Let "$A \circledast B$" be defined as $A^B + A \cdot B$. Calculate $4 \circledast 3$.
2. Solve this system of two linear equations:

$$\begin{aligned} 2x - y &= 4 \\ x + y &= 5 \end{aligned}$$

Functions

Introduction to Functions

- A **function** is a relation (or mapping or transformation) that assigns every member of the domain to exactly one member of the range.

Introduction to Functions

- A **function** is a relation (or mapping or transformation) that assigns every member of the domain to exactly one member of the range.
- Domain: set of all possible inputs. Set X over which $f(x)$ defined.

Introduction to Functions

- A **function** is a relation (or mapping or transformation) that assigns every member of the domain to exactly one member of the range.
- Domain: set of all possible inputs. Set X over which $f(x)$ defined.
- Range ("Image"): set of all possible (obtained) output values. Set of elements of Y assigned to elements of X by $f(x)$. $f(X) = \{y : y = f(x), x \in X\}$

Introduction to Functions

- A **function** is a relation (or mapping or transformation) that assigns every member of the domain to exactly one member of the range.
- Domain: set of all possible inputs. Set X over which $f(x)$ defined.
- Range ("Image"): set of all possible (obtained) output values. Set of elements of Y assigned to elements of X by $f(x)$. $f(X) = \{y : y = f(x), x \in X\}$
- Mapping notation examples

Introduction to Functions

- A **function** is a relation (or mapping or transformation) that assigns every member of the domain to exactly one member of the range.
- Domain: set of all possible inputs. Set X over which $f(x)$ defined.
- Range ("Image"): set of all possible (obtained) output values. Set of elements of Y assigned to elements of X by $f(x)$. $f(X) = \{y : y = f(x), x \in X\}$
- Mapping notation examples
  - $f : X \to Y$

Introduction to Functions

- A **function** is a relation (or mapping or transformation) that assigns every member of the domain to exactly one member of the range.
- Domain: set of all possible inputs. Set X over which $f(x)$ defined.
- Range ("Image"): set of all possible (obtained) output values. Set of elements of Y assigned to elements of X by $f(x)$. $f(X) = \{y : y = f(x), x \in X\}$
- Mapping notation examples
  - $f : X \to Y$
  - Function of one variable: $f : \mathbb{R}^1 \to \mathbb{R}^1$

Introduction to Functions

- A **function** is a relation (or mapping or transformation) that assigns every member of the domain to exactly one member of the range.
- Domain: set of all possible inputs. Set X over which $f(x)$ defined.
- Range ("Image"): set of all possible (obtained) output values. Set of elements of Y assigned to elements of X by $f(x)$. $f(X) = \{y : y = f(x), x \in X\}$
- Mapping notation examples
  - $f : X \to Y$
  - Function of one variable: $f : \mathbb{R}^1 \to \mathbb{R}^1$
  - Function of two variables: $f : \mathbb{R}^2 \to \mathbb{R}^1$, $f : \mathbb{R}^2 \to \mathbb{R}^2$

Functions

Examples:

- $f(x) = x + 1$

For each x in $\mathbb{R}^1$, $f(x)$ assigns the number $x + 1$.

Functions

Examples:

- $f(x) = x + 1$

For each x in $\mathbb{R}^1$, $f(x)$ assigns the number $x + 1$.

- $f(x, y) = x^2 + y^2$

For each ordered pair (x, y) in $\mathbb{R}^2$, $f(x, y)$ assigns the number $x^2 + y^2$

Functions

Examples:

- $f(x) = x + 1$

For each x in $\mathbb{R}^1$, $f(x)$ assigns the number $x + 1$.

- $f(x, y) = x^2 + y^2$

For each ordered pair (x, y) in $\mathbb{R}^2$, $f(x, y)$ assigns the number $x^2 + y^2$

- Often use one variable x as input and another y as output.

Example: $y = x + 1$

Functions

Examples:

- $f(x) = x + 1$

For each x in $\mathbb{R}^1$, $f(x)$ assigns the number $x + 1$.

- $f(x, y) = x^2 + y^2$

For each ordered pair (x, y) in $\mathbb{R}^2$, $f(x, y)$ assigns the number $x^2 + y^2$

- Often use one variable x as input and another y as output.

Example: $y = x + 1$

- Input variable: *predictor*, *covariate*, *indep var*
- Output variable: *outcome*, *response*, *dep var*

Example 1

$f(x) = \frac{3}{1+x^2}$

Domain $X =$

Example 1

$f(x) = \frac{3}{1+x^2}$

Domain $X = \mathbb{R}^1$

Example 1

$f(x) = \frac{3}{1+x^2}$

Domain $X = \mathbb{R}^1$

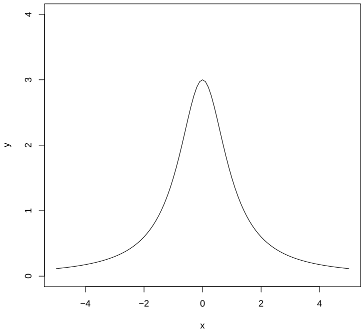
Range $f(X) =$

Example 1

$f(x) = \frac{3}{1+x^2}$

Domain $X = \mathbb{R}^1$

Range $f(X) = (0, 3]$



Example 2

$$f(x) = \begin{cases} x+1, & 1 \le x \le 2 \\ 0, & x = 0 \\ 1-x & -2 \le x \le -1 \end{cases}$$

Domain $X =$

Example 2

$$f(x) = \begin{cases} x+1, & 1 \le x \le 2 \\ 0, & x = 0 \\ 1-x & -2 \le x \le -1 \end{cases}$$

Domain $X = [-2,-1] \cup \{0\} \cup [1,2]$

Example 2

$$f(x) = \begin{cases} x+1, & 1 \leq x \leq 2 \\ 0, & x = 0 \\ 1-x & -2 \leq x \leq -1 \end{cases}$$

Domain $X = [-2,-1] \cup \{0\} \cup [1,2]$

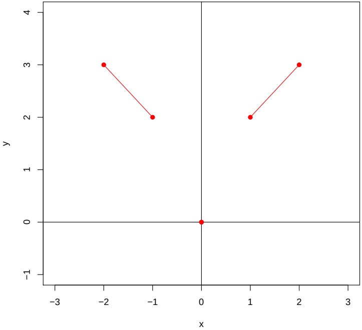
Range $f(X) =$

Example 2

$$f(x) = \begin{cases} x+1, & 1 \le x \le 2 \\ 0, & x = 0 \\ 1-x & -2 \le x \le -1 \end{cases}$$

Domain $X = [-2, -1] \cup \{0\} \cup [1, 2]$

Range $f(X) = [2, 3] \cup \{0\}$



Example 3

$f(x) = \frac{1}{x}$

Domain $X =$

Example 3

$f(x) = \frac{1}{x}$

Domain $X = \mathbb{R}^1 \setminus \{0\}$

Example 3

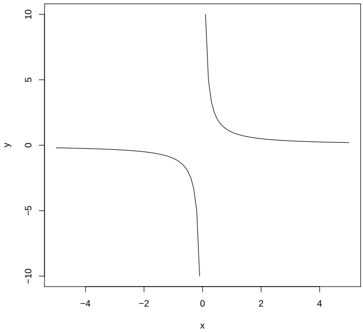
$f(x) = \frac{1}{x}$

Domain $X = \mathbb{R}^1 \setminus \{0\}$
Range $f(X) =$

Example 3

$f(x) = \frac{1}{x}$

Domain $X = \mathbb{R}^1 \setminus \{0\}$
Range $f(X) = \mathbb{R}^1 \setminus \{0\}$



Example 4

$f(x, y) = x^2 + y^2$

Domain $X = \mathbb{R}^2$

Example 4

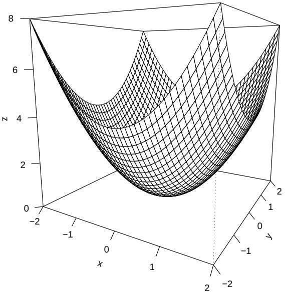
$f(x, y) = x^2 + y^2$

Domain $X = \mathbb{R}^2$
Image $f(X, Y) =$

Example 4

$$f(x, y) = x^2 + y^2$$

Domain $X = \mathbb{R}^2$
Image $f(X, Y) = \mathbb{R}^1_+ \cup \{0\}$



General Types of Functions

- **Monomials**: $f(x) = ax^k$

a is the coefficient. k is the degree.

Examples: $y = x^2$, $y = -\frac{1}{2}x^3$

General Types of Functions

- **Monomials**: $f(x) = ax^k$

a is the coefficient. k is the degree.

Examples: $y = x^2$, $y = -\frac{1}{2}x^3$

- **Polynomials**: sum of monomials.

Examples: $y = -\frac{1}{2}x^3 + x^2$, $y = 3x + 5$

The degree of a polynomial is the highest degree of its monomial terms. Write polynomials with terms in decreasing degree.

General Types of Functions

- **Monomials**: $f(x) = ax^k$

a is the coefficient. k is the degree.

Examples: $y = x^2$, $y = -\frac{1}{2}x^3$

- **Polynomials**: sum of monomials.

Examples: $y = -\frac{1}{2}x^3 + x^2$, $y = 3x + 5$

The degree of a polynomial is the highest degree of its monomial terms. Write polynomials with terms in decreasing degree.

- **Rational Functions**: ratio of two polynomials.

Examples: $y = \frac{x}{2}$, $y = \frac{x^2+1}{x^2-2x+1}$

- **Exponential Functions**: Example: $y = 2^x$

General Types of Functions

- **Monomials**: $f(x) = ax^k$

a is the coefficient. k is the degree.

Examples: $y = x^2$, $y = -\frac{1}{2}x^3$

- **Polynomials**: sum of monomials.

Examples: $y = -\frac{1}{2}x^3 + x^2$, $y = 3x + 5$

The degree of a polynomial is the highest degree of its monomial terms. Write polynomials with terms in decreasing degree.

- **Rational Functions**: ratio of two polynomials.

Examples: $y = \frac{x}{2}$, $y = \frac{x^2+1}{x^2-2x+1}$

- **Exponential Functions**: Example: $y = 2^x$
- **Trigonometric Functions**: Examples: $y = \cos(x)$, $y = 3\sin(4x)$

Trigonometric Functions: Gill & Casella (2004)

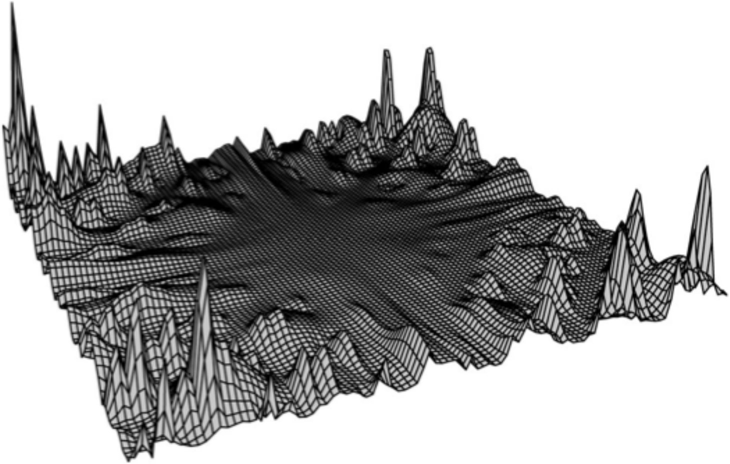


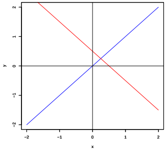
Fig. 1 A highly multimodal surface.

$$\begin{aligned} f(x,y) &= |(x\sin(20y-90) - y\cos(20x+45))^3 a\cos(\sin(90y+42)x) \\ &\quad +(x\cos(10y+10) - y\sin(10x+15))^2 a\cos(\cos(10x+24)y)| \end{aligned}$$

Linear and Nonlinear Functions

- **Linear**: polynomial of degree 1.

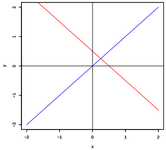
Example: $y = mx + b$, where $m =$ slope and $b = y$-intercept.



Linear and Nonlinear Functions

- **Linear**: polynomial of degree 1.

Example: $y = mx + b$, where $m =$ slope and $b = y$-intercept.



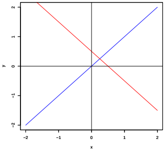
- **Nonlinear**: anything that isn't constant or polynomial of degree 1.

Examples: $y = x^2 + 2x + 1$, $y = \sin(x)$, $y = \ln(x)$, $y = e^x$

Linear and Nonlinear Functions

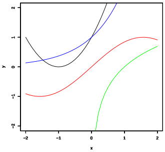
- **Linear**: polynomial of degree 1.

Example: $y = mx + b$, where m = slope and b = y-intercept.



- **Nonlinear**: anything that isn't constant or polynomial of degree 1.

Examples: $y = x^2 + 2x + 1$, $y = \sin(x)$, $y = \ln(x)$, $y = e^x$



Exponents and Logarithms

Definition:

$$(\log_b P = Q) \iff (b^Q = P)$$

Exponents and Logarithms

Definition:

$$(\log_b P = Q) \iff (b^Q = P)$$

Important properties of logarithms:

1. $b^{\log_b x} = x$
2. $\log_b(xy) = \log_b(x) + \log_b(y)$
3. $\log_b(\frac{x}{y}) = \log_b(x) - \log_b(y)$
4. $\log_b(y^x) = x \log_b y$
5. $\log_b(\frac{1}{x}) = -\log_b(x)$

Exponents and Logarithms

Definition:

$$(\log_b P = Q) \iff (b^Q = P)$$

Important properties of logarithms:

1. $b^{\log_b x} = x$
2. $\log_b(xy) = \log_b(x) + \log_b(y)$
3. $\log_b(\frac{x}{y}) = \log_b(x) - \log_b(y)$
4. $\log_b(y^x) = x \log_b y$
5. $\log_b(\frac{1}{x}) = -\log_b(x) \qquad b^{-x} = \frac{1}{b^x}$

Exponents and Logarithms

Definition:

$$(\log_b P = Q) \iff (b^Q = P)$$

Important properties of logarithms:

1. $b^{\log_b x} = x$
2. $\log_b(xy) = \log_b(x) + \log_b(y)$
3. $\log_b(\frac{x}{y}) = \log_b(x) - \log_b(y)$
4. $\log_b(y^x) = x \log_b y$
5. $\log_b(\frac{1}{x}) = -\log_b(x) \qquad b^{-x} = \frac{1}{b^x}$
6. $\log_b(1) = 0$

Exponents and Logarithms

Definition:

$$(\log_b P = Q) \iff (b^Q = P)$$

Important properties of logarithms:

1. $b^{\log_b x} = x$
2. $\log_b(xy) = \log_b(x) + \log_b(y)$
3. $\log_b(\frac{x}{y}) = \log_b(x) - \log_b(y)$
4. $\log_b(y^x) = x \log_b y$
5. $\log_b(\frac{1}{x}) = -\log_b(x) \qquad b^{-x} = \frac{1}{b^x}$
6. $\log_b(1) = 0 \qquad b^0 = 1$

Exponents and Logarithms

Definition:

$$(\log_b P = Q) \iff (b^Q = P)$$

Important properties of logarithms:

1. $b^{\log_b x} = x$
2. $\log_b(xy) = \log_b(x) + \log_b(y)$
3. $\log_b(\frac{x}{y}) = \log_b(x) - \log_b(y)$
4. $\log_b(y^x) = x \log_b y$
5. $\log_b(\frac{1}{x}) = -\log_b(x) \qquad b^{-x} = \frac{1}{b^x}$
6. $\log_b(1) = 0 \qquad b^0 = 1$
7. $\log_b(x) = \frac{\log_a(x)}{\log_a(b)}$

Examples

- $\log(\sqrt{10}) =$

Examples

- $\log(\sqrt{10}) = \frac{1}{2}$

Examples

- $\log(\sqrt{10}) = \frac{1}{2}$
- $\log(1) =$

Examples

- $\log(\sqrt{10}) = \frac{1}{2}$
- $\log(1) = 0$

Examples

- $\log(\sqrt{10}) = \frac{1}{2}$
- $\log(1) = 0$
- $\log(10) =$

Examples

- $\log(\sqrt{10}) = \frac{1}{2}$
- $\log(1) = 0$
- $\log(10) = 1$

Examples

- $\log(\sqrt{10}) = \frac{1}{2}$
- $\log(1) = 0$
- $\log(10) = 1$
- $\log(100) =$

Examples

- $\log(\sqrt{10}) = \frac{1}{2}$
- $\log(1) = 0$
- $\log(10) = 1$
- $\log(100) = 2$

Examples

- $\log(\sqrt{10}) = \frac{1}{2}$
- $\log(1) = 0$
- $\log(10) = 1$
- $\log(100) = 2$
- $\ln(1) =$

Examples

- $\log(\sqrt{10}) = \frac{1}{2}$
- $\log(1) = 0$
- $\log(10) = 1$
- $\log(100) = 2$
- $\ln(1) = 0$

Examples

- $\log(\sqrt{10}) = \ \frac{1}{2}$
- $\log(1) = \ 0$
- $\log(10) = \ 1$
- $\log(100) = \ 2$
- $\ln(1) = \ 0$
- $\ln(e) =$

Examples

- $\log(\sqrt{10}) = \frac{1}{2}$
- $\log(1) = 0$
- $\log(10) = 1$
- $\log(100) = 2$
- $\ln(1) = 0$
- $\ln(e) = 1$

Examples

- $\log(\sqrt{10}) = \frac{1}{2}$
- $\log(1) = 0$
- $\log(10) = 1$
- $\log(100) = 2$
- $\ln(1) = 0$
- $\ln(e) = 1$

Summation Notation

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + \cdots + x_n$$

Summation Notation

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + \cdots + x_n$$

Let $x = (4, -2, 3)$, $y = (-1, 0, 1)$. Calculate

- $\sum_{i=1}^{2} x_i$
- $\sum_{i=1}^{3} x_i^{y_i}$
- $\sum_{i=2}^{3} (x_i + y_{i-1})$

Summation Notation

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + \cdots + x_n$$

Let $x = (4, -2, 3)$, $y = (-1, 0, 1)$. Calculate

- $\sum_{i=1}^{2} x_i$
- $\sum_{i=1}^{3} x_i^{y_i}$
- $\sum_{i=2}^{3} (x_i + y_{i-1})$

- $\sum_{i=1}^{2} x_i = 4 + (-2) = 2$

Summation Notation

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + \cdots + x_n$$

Let $x = (4, -2, 3)$, $y = (-1, 0, 1)$. Calculate

- $\sum_{i=1}^{2} x_i$
- $\sum_{i=1}^{3} x_i^{y_i}$
- $\sum_{i=2}^{3} (x_i + y_{i-1})$

- $\sum_{i=1}^{2} x_i = 4 + (-2) = 2$
- $\sum_{i=1}^{3} x_i^{y_i} = 4^{-1} + (-2)^0 + 3^1 = \frac{1}{4} + 1 + 3 = 4\frac{1}{4}$

Summation Notation

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + \cdots + x_n$$

Let $x = (4, -2, 3)$, $y = (-1, 0, 1)$. Calculate

- $\sum_{i=1}^{2} x_i$
- $\sum_{i=1}^{3} x_i^{y_i}$
- $\sum_{i=2}^{3} (x_i + y_{i-1})$

- $\sum_{i=1}^{2} x_i = 4 + (-2) = 2$
- $\sum_{i=1}^{3} x_i^{y_i} = 4^{-1} + (-2)^0 + 3^1 = \frac{1}{4} + 1 + 3 = 4\frac{1}{4}$
- $\sum_{i=2}^{3} (x_i + y_{i-1}) = (-2 + -1) + (3 + 0) = 0$

Summation Notation

Properties:

- $\sum_{i=1}^{n} cx_i = c\sum_{i=1}^{n} x_i$
- $\sum_{i=1}^{n} (x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$
- $\sum_{i=1}^{n} c = nc$

Product Notation

$$\prod_{i=1}^{n} x_i = x_1 x_2 x_3 \cdots x_n$$

Product Notation

$$\prod_{i=1}^{n} x_i = x_1 x_2 x_3 \cdots x_n$$

Let $x = (4, -2, 3)$, $y = (-1, 0, 1)$. Calculate

- $\prod_{i=1}^{2} x_i$
- $\prod_{i=1}^{3} x_i^{y_i}$
- $\prod_{i=2}^{3} (x_i + y_{i-1})$

Product Notation

$$\prod_{i=1}^{n} x_i = x_1 x_2 x_3 \cdots x_n$$

Let $x = (4, -2, 3)$, $y = (-1, 0, 1)$. Calculate

- $\prod_{i=1}^{2} x_i$
- $\prod_{i=1}^{3} x_i^{y_i}$
- $\prod_{i=2}^{3} (x_i + y_{i-1})$

- $\prod_{i=1}^{2} x_i = 4 \cdot -2 = -8$

Product Notation

$$\prod_{i=1}^{n} x_i = x_1 x_2 x_3 \cdots x_n$$

Let $x = (4, -2, 3)$, $y = (-1, 0, 1)$. Calculate

- $\prod_{i=1}^{2} x_i$
- $\prod_{i=1}^{3} x_i^{y_i}$
- $\prod_{i=2}^{3} (x_i + y_{i-1})$

- $\prod_{i=1}^{2} x_i = 4 \cdot -2 = -8$
- $\prod_{i=1}^{3} x_i^{y_i} = 4^{-1} \cdot (-2)^0 \cdot 3^1 = \frac{1}{4} \cdot 1 \cdot 3 = \frac{3}{4}$

Product Notation

$$\prod_{i=1}^{n} x_i = x_1 x_2 x_3 \cdots x_n$$

Let $x = (4, -2, 3)$, $y = (-1, 0, 1)$. Calculate

- $\prod_{i=1}^{2} x_i$
- $\prod_{i=1}^{3} x_i^{y_i}$
- $\prod_{i=2}^{3} (x_i + y_{i-1})$

- $\prod_{i=1}^{2} x_i = 4 \cdot -2 = -8$
- $\prod_{i=1}^{3} x_i^{y_i} = 4^{-1} \cdot (-2)^0 \cdot 3^1 = \frac{1}{4} \cdot 1 \cdot 3 = \frac{3}{4}$
- $\prod_{i=2}^{3} (x_i + y_{i-1}) = (-2 + -1) \cdot (3 + 0) = -9$

Product Notation

Properties:

- $\prod_{i=1}^{n} cx_i = c^n \prod_{i=1}^{n} x_i$
- $\prod_{i=1}^{n} (x_i + y_i) = (x_1 + y_1)(x_2 + y_2) \dots$
- $\prod_{i=1}^{n} c = c^n$

Sums, Products, and Logs

Logarithms can help us change products into sums.

$$\log\left(\prod_{i=1}^{n} cx_i\right)$$

Sums, Products, and Logs

Logarithms can help us change products into sums.

$$\log\left(\prod_{i=1}^{n} cx_i\right) \quad = \quad \log(cx_1 \cdot cx_2 \cdot \ldots \cdot cx_n)$$

Sums, Products, and Logs

Logarithms can help us change products into sums.

$$\begin{aligned}
\log\left(\prod_{i=1}^{n} cx_i\right) &= \log(cx_1 \cdot cx_2 \cdot \ldots \cdot cx_n) \\
&= \log(cx_1) + \log(cx_2) + \ldots + \log(cx_n)
\end{aligned}$$

Sums, Products, and Logs

Logarithms can help us change products into sums.

$$\begin{aligned}
\log\left(\prod_{i=1}^{n} cx_i\right) &= \log(cx_1 \cdot cx_2 \cdot \ldots \cdot cx_n) \\
&= \log(cx_1) + \log(cx_2) + \ldots + \log(cx_n) \\
&= \sum_{i=1}^{n} \log(cx_i)
\end{aligned}$$

Sums, Products, and Logs

Logarithms can help us change products into sums.

$$
\begin{aligned}
\log\left(\prod_{i=1}^{n} cx_i\right) &= \log(cx_1 \cdot cx_2 \cdot \ldots \cdot cx_n) \\
&= \log(cx_1) + \log(cx_2) + \ldots + \log(cx_n) \\
&= \sum_{i=1}^{n} \log(cx_i) \\
&= [(\log c + \log x_1) + (\log c + \log x_2) + \\
&\qquad \ldots + (\log c + \log x_n)]
\end{aligned}
$$

Sums, Products, and Logs

Logarithms can help us change products into sums.

$$
\begin{aligned}
\log\left(\prod_{i=1}^{n} cx_i\right) &= \log(cx_1 \cdot cx_2 \cdot \ldots \cdot cx_n) \\
&= \log(cx_1) + \log(cx_2) + \ldots + \log(cx_n) \\
&= \sum_{i=1}^{n} \log(cx_i) \\
&= [(\log c + \log x_1) + (\log c + \log x_2) + \\
&\quad \ldots + (\log c + \log x_n)] \\
&= \sum_{i=1}^{n} \log c + \sum_{i=1}^{n} \log(x_i)
\end{aligned}
$$

Sums, Products, and Logs

Logarithms can help us change products into sums.

$$
\begin{aligned}
\log\left(\prod_{i=1}^{n} cx_i\right) &= \log(cx_1 \cdot cx_2 \cdot \ldots \cdot cx_n) \\
&= \log(cx_1) + \log(cx_2) + \ldots + \log(cx_n) \\
&= \sum_{i=1}^{n} \log(cx_i) \\
&= [(\log c + \log x_1) + (\log c + \log x_2) + \\
&\quad \ldots + (\log c + \log x_n)] \\
&= \sum_{i=1}^{n} \log c + \sum_{i=1}^{n} \log(x_i) \\
&= n \log(c) + \sum_{i=1}^{n} \log(x_i)
\end{aligned}
$$

Vectors

Vectors

- **Vector**: A vector in n-space is an ordered list of n numbers. These numbers can be represented as either a row vector or a column vector:

$$\mathbf{v} = \begin{pmatrix} v_1 & v_2 & \dots & v_n \end{pmatrix}, \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

Vectors

- **Vector**: A vector in n-space is an ordered list of n numbers. These numbers can be represented as either a row vector or a column vector:

$$\mathbf{v} = \begin{pmatrix} v_1 & v_2 & \dots & v_n \end{pmatrix}, \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

- We can also think of a vector as defining a point in n-dimensional space, usually $\mathbf{R}^n$; each element of the vector defines the coordinate of the point in a particular direction.

Vector Arithmetic

- **Vector Addition**: Vector addition is defined for two vectors $\mathbf{u}$ and $\mathbf{v}$ iff they have the same number of elements:

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 + v_1 & u_2 + v_2 & \cdots & u_n + v_n \end{pmatrix}$$

Vector Arithmetic

- **Vector Addition**: Vector addition is defined for two vectors $\mathbf{u}$ and $\mathbf{v}$ iff they have the same number of elements:

$$\mathbf{u}+\mathbf{v}=\begin{pmatrix} u_1+v_1 & u_2+v_2 & \cdots & u_n+v_n \end{pmatrix}$$

Let $\mathbf{u}=\begin{pmatrix} 3 & -2 & 1 \end{pmatrix}$, $\mathbf{v}=\begin{pmatrix} 2 & 0 & 1 \end{pmatrix}$.

Vector Arithmetic

- **Vector Addition**: Vector addition is defined for two vectors $\mathbf{u}$ and $\mathbf{v}$ iff they have the same number of elements:

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 + v_1 & u_2 + v_2 & \cdots & u_n + v_n \end{pmatrix}$$

Let $\mathbf{u} = \begin{pmatrix} 3 & -2 & 1 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 2 & 0 & 1 \end{pmatrix}$.

$\mathbf{u} + \mathbf{v} =$

Vector Arithmetic

- **Vector Addition**: Vector addition is defined for two vectors $\mathbf{u}$ and $\mathbf{v}$ iff they have the same number of elements:

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 + v_1 & u_2 + v_2 & \cdots & u_n + v_n \end{pmatrix}$$

Let $\mathbf{u} = \begin{pmatrix} 3 & -2 & 1 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 2 & 0 & 1 \end{pmatrix}$.

$\mathbf{u} + \mathbf{v} = \begin{pmatrix} 3 + 2 & -2 + 0 & 1 + 1 \end{pmatrix}$

Vector Arithmetic

- **Vector Addition**: Vector addition is defined for two vectors $\mathbf{u}$ and $\mathbf{v}$ iff they have the same number of elements:

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 + v_1 & u_2 + v_2 & \cdots & u_n + v_n \end{pmatrix}$$

Let $\mathbf{u} = \begin{pmatrix} 3 & -2 & 1 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 2 & 0 & 1 \end{pmatrix}$.

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} 3+2 & -2+0 & 1+1 \end{pmatrix} = \begin{pmatrix} 5 & -2 & 2 \end{pmatrix}$$

```r
u <- c(3, -2, 1)
v <- c(2, 0, 1)
u + v
```

```
## [1]  5 -2  2
```

Vector Arithmetic

- **Scalar Multiplication**: The product of a scalar c and vector $\mathbf{v}$ is:
$$c\mathbf{v} = \begin{pmatrix} cv_1 & cv_2 & \dots & cv_n \end{pmatrix}$$

Vector Arithmetic

- **Scalar Multiplication**: The product of a scalar c and vector $\mathbf{v}$ is:
$$c\mathbf{v} = \begin{pmatrix} cv_1 & cv_2 & \dots & cv_n \end{pmatrix}$$

Let $\mathbf{v} = \begin{pmatrix} 3 & -2 & 1 \end{pmatrix}$, $c = 6$.

Vector Arithmetic

- **Scalar Multiplication**: The product of a scalar c and vector $\mathbf{v}$ is:
$$c\mathbf{v} = \begin{pmatrix} cv_1 & cv_2 & \dots & cv_n \end{pmatrix}$$

Let $\mathbf{v} = \begin{pmatrix} 3 & -2 & 1 \end{pmatrix}$, $c = 6$.

$c\mathbf{v} =$

Vector Arithmetic

- **Scalar Multiplication**: The product of a scalar c and vector $\mathbf{v}$ is:
$$c\mathbf{v} = \begin{pmatrix} cv_1 & cv_2 & \dots & cv_n \end{pmatrix}$$

Let $\mathbf{v} = \begin{pmatrix} 3 & -2 & 1 \end{pmatrix}$, $c = 6$.

$c\mathbf{v} = \begin{pmatrix} 6 \cdot 3 & 6 \cdot -2 & 6 \cdot 1 \end{pmatrix}$

Vector Arithmetic

- **Scalar Multiplication**: The product of a scalar c and vector $\mathbf{v}$ is:
$$c\mathbf{v} = \begin{pmatrix} cv_1 & cv_2 & \dots & cv_n \end{pmatrix}$$

Let $\mathbf{v} = \begin{pmatrix} 3 & -2 & 1 \end{pmatrix}$, $c = 6$.

$$c\mathbf{v} = \begin{pmatrix} 6 \cdot 3 & 6 \cdot -2 & 6 \cdot 1 \end{pmatrix} = \begin{pmatrix} 18 & -12 & 6 \end{pmatrix}$$

```r
c <- 6
v <- c(3, -2, 1)
c * v
```

```
## [1]  18 -12   6
```

Vector Properties

Let $\mathbf{u}, \mathbf{v}, \mathbf{w}$ be vectors, c, d be constants, $\mathbf{0}$ be vector $\mathbf{z}$ s.t. $z_i = 0$.

Vector Properties

Let $\mathbf{u}, \mathbf{v}, \mathbf{w}$ be vectors, c, d be constants, $\mathbf{0}$ be vector $\mathbf{z}$ s.t. $z_i = 0$.

Then, assuming $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{0}$ *conformable*,

Vector Properties

Let $\mathbf{u}, \mathbf{v}, \mathbf{w}$ be vectors, c, d be constants, $\mathbf{0}$ be vector $\mathbf{z}$ s.t. $z_i = 0$.

Then, assuming $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{0}$ *conformable*,

- Commutativity: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
- Associativity: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
- Distributivity: $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
- Scalar Distributivity: $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$

Vector Properties

Let $\mathbf{u}, \mathbf{v}, \mathbf{w}$ be vectors, c, d be constants, $\mathbf{0}$ be vector $\mathbf{z}$ s.t. $z_i = 0$.

Then, assuming $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{0}$ *conformable*,

- Commutativity: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
- Associativity: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
- Distributivity: $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
- Scalar Distributivity: $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$
- Additive Identity: $\mathbf{u} + \mathbf{0} = \mathbf{u}$

Vector Properties

Let $\mathbf{u}, \mathbf{v}, \mathbf{w}$ be vectors, c, d be constants, $\mathbf{0}$ be vector $\mathbf{z}$ s.t. $z_i = 0$.

Then, assuming $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{0}$ *conformable*,

- Commutativity: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
- Associativity: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
- Distributivity: $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
- Scalar Distributivity: $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$
- Additive Identity: $\mathbf{u} + \mathbf{0} = \mathbf{u}$
- Scalar Multiplicative Identity: $1\mathbf{u} = \mathbf{u}$

Transpose

- **Transpose**: The transpose of a vector $\mathbf{u}$ changes $\mathbf{u}$'s row/column status.

Transpose

- **Transpose**: The transpose of a vector $\mathbf{u}$ changes $\mathbf{u}$'s row/column status.
$$\mathbf{u}^T = \mathbf{u}'$$

Transpose

- **Transpose**: The transpose of a vector $\mathbf{u}$ changes $\mathbf{u}$'s row/column status.
$$\mathbf{u}^T = \mathbf{u}'$$

Let $\mathbf{u} = \begin{pmatrix} 3 & -2 & 1 \end{pmatrix}$ be 1×3.

Transpose

- **Transpose**: The transpose of a vector $\mathbf{u}$ changes $\mathbf{u}$'s row/column status.
$$\mathbf{u}^T = \mathbf{u}'$$

Let $\mathbf{u} = \begin{pmatrix} 3 & -2 & 1 \end{pmatrix}$ be 1×3. Then, $\mathbf{u}' = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$ is 3×1.

Inner Product

- **Inner Product**: The Euclidean inner product (also, the "dot product") of two vectors $\mathbf{u}$ and $\mathbf{v}$ is defined iff they have the same number of elements

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i$$

Inner Product

- **Inner Product**: The Euclidean inner product (also, the "dot product") of two vectors $\mathbf{u}$ and $\mathbf{v}$ is defined iff they have the same number of elements

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i$$

Let $\mathbf{u} = \begin{pmatrix} 3 & -2 & 1 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 2 & 0 & 1 \end{pmatrix}$.

Inner Product

- **Inner Product**: The Euclidean inner product (also, the "dot product") of two vectors $\mathbf{u}$ and $\mathbf{v}$ is defined iff they have the same number of elements

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i$$

Let $\mathbf{u} = \begin{pmatrix} 3 & -2 & 1 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 2 & 0 & 1 \end{pmatrix}$.

$\mathbf{u} \cdot \mathbf{v} =$

Inner Product

- **Inner Product**: The Euclidean inner product (also, the "dot product") of two vectors $\mathbf{u}$ and $\mathbf{v}$ is defined iff they have the same number of elements

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i$$

Let $\mathbf{u} = \begin{pmatrix} 3 & -2 & 1 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 2 & 0 & 1 \end{pmatrix}$.

$\mathbf{u} \cdot \mathbf{v} = 3 \cdot 2 + -2 \cdot 0 + 1 \cdot 1$

Inner Product

- **Inner Product**: The Euclidean inner product (also, the "dot product") of two vectors $\mathbf{u}$ and $\mathbf{v}$ is defined iff they have the same number of elements

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i$$

Let $\mathbf{u} = \begin{pmatrix} 3 & -2 & 1 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 2 & 0 & 1 \end{pmatrix}$.

$\mathbf{u} \cdot \mathbf{v} = 3 \cdot 2 + -2 \cdot 0 + 1 \cdot 1 = 6 + 1 = 7.$

```
u %*% v |> drop() # Simplify array dims
```

```
## [1] 7
```

If $\mathbf{u} \cdot \mathbf{v} = 0$, the vectors are *orthogonal* (or perpendicular).

Let $\mathbf{u} = \begin{pmatrix} 5 & 0 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 0 & -2 \end{pmatrix}$.

Inner Product and Orthogonality

If $\mathbf{u} \cdot \mathbf{v} = 0$, the vectors are *orthogonal* (or perpendicular).

Let $\mathbf{u} = \begin{pmatrix} 5 & 0 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 0 & -2 \end{pmatrix}$.

$\mathbf{u} \cdot \mathbf{v} = 5 \cdot 0 + 0 \cdot -2 = 0 + 0 = 0.$

Inner Product

- Think about the $\mathbf{u} \cdot \mathbf{v}$ inner product as

$$\underbrace{\mathbf{u}}_{1\times k} \cdot \underbrace{\mathbf{v}}_{k\times 1} = \underbrace{w}_{1\times 1}$$

Inner Product

- Think about the $\mathbf{u} \cdot \mathbf{v}$ inner product as

$$\underbrace{\mathbf{u}}_{1\times k} \cdot \underbrace{\mathbf{v}}_{k\times 1} = \underbrace{w}_{1\times 1}$$

- Or, assume $\mathbf{u}$, $\mathbf{v}$ both $k \times 1$ columns, then $\mathbf{u} \cdot \mathbf{v} \equiv \mathbf{v}^{\mathbf{T}}\mathbf{u}$

Inner Product

- Think about the $\mathbf{u} \cdot \mathbf{v}$ inner product as

$$\underbrace{\mathbf{u}}_{1\times k} \cdot \underbrace{\mathbf{v}}_{k\times 1} = \underbrace{w}_{1\times 1}$$

- Or, assume $\mathbf{u}$, $\mathbf{v}$ both $k \times 1$ columns, then $\mathbf{u} \cdot \mathbf{v} \equiv \mathbf{v}^{\mathbf{T}}\mathbf{u}$

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v}'\mathbf{u} = \begin{pmatrix} 2 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$$

Inner Product

- Think about the $\mathbf{u} \cdot \mathbf{v}$ inner product as

$$\underbrace{\mathbf{u}}_{1 \times k} \cdot \underbrace{\mathbf{v}}_{k \times 1} = \underbrace{w}_{1 \times 1}$$

- Or, assume $\mathbf{u}$, $\mathbf{v}$ both $k \times 1$ columns, then $\mathbf{u} \cdot \mathbf{v} \equiv \mathbf{v^T u}$

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v}'\mathbf{u} = \begin{pmatrix} 2 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} = 7$$

Inner Product Properties

- Commutativity: $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
- Associativity: $c(\mathbf{u} \cdot \mathbf{v}) = (c\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (c\mathbf{v})$
- Distributivity: $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
- Zero Product: $\mathbf{u} \cdot \mathbf{0} = 0$

Vector Norm

- **Vector Norm**: The *norm* of a vector measures its length. There are many norms; most common: Euclidean norm (corresponds to usual conception of distance in 3D space):

Vector Norm

- **Vector Norm**: The *norm* of a vector measures its length. There are many norms; most common: Euclidean norm (corresponds to usual conception of distance in 3D space):

$$||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1 v_1 + v_2 v_2 + \cdots + v_n v_n}$$

Vector Norm

- **Vector Norm**: The *norm* of a vector measures its length. There are many norms; most common: Euclidean norm (corresponds to usual conception of distance in 3D space):

$$||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1 v_1 + v_2 v_2 + \cdots + v_n v_n}$$

$$||\mathbf{v}|| \quad = \quad || \begin{pmatrix} 2 & 0 & 1 \end{pmatrix} ||$$

Vector Norm

- **Vector Norm**: The *norm* of a vector measures its length. There are many norms; most common: Euclidean norm (corresponds to usual conception of distance in 3D space):

$$||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1 v_1 + v_2 v_2 + \cdots + v_n v_n}$$

$$\begin{aligned} ||\mathbf{v}|| &= || \begin{pmatrix} 2 & 0 & 1 \end{pmatrix} || \\ &= \sqrt{2^2 + 0^2 + 1^2} \end{aligned}$$

Vector Norm

- **Vector Norm**: The *norm* of a vector measures its length. There are many norms; most common: Euclidean norm (corresponds to usual conception of distance in 3D space):

$$||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1 v_1 + v_2 v_2 + \cdots + v_n v_n}$$

$$\begin{aligned} ||\mathbf{v}|| &= || \begin{pmatrix} 2 & 0 & 1 \end{pmatrix} || \\ &= \sqrt{2^2 + 0^2 + 1^2} \\ &= \sqrt{5} \end{aligned}$$

Vector Norm Properties

- Scalar Multiplication: $||c\mathbf{u}|| = |c| \cdot ||\mathbf{u}||$
- Vector Distance: $d(\mathbf{u}, \mathbf{v}) = ||\mathbf{u} - \mathbf{v}||$
- Norm Squared: $||\mathbf{u}||^2 = \mathbf{u} \cdot \mathbf{u}$
- Cosine Rule: $\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{||\mathbf{u}|| \cdot ||\mathbf{v}||}$
- Difference Norm: $||\mathbf{u} - \mathbf{v}||^2 = ||\mathbf{u}||^2 - 2(\mathbf{u} \cdot \mathbf{v}) + ||\mathbf{v}||^2$
  $= ||\mathbf{u}||^2 - 2||\mathbf{u}|| \cdot ||\mathbf{v}||(\cos\theta) + ||\mathbf{v}||^2$
- Triangle Inequality: $||\mathbf{u} + \mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||$
- Cauchy-Schwartz Inequality: $||\mathbf{u} \cdot \mathbf{v}|| \leq ||\mathbf{u}|| \cdot ||\mathbf{v}||$

Mahalanobis Distance

$$MD_{ij} = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)'\Sigma^{-1}(\mathbf{x}_i - \mathbf{x}_j)}$$

Mahalanobis Distance

$$MD_{ij} = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)'\Sigma^{-1}(\mathbf{x}_i - \mathbf{x}_j)}$$

Like Euclidean distance, but scaled by inverse covariances

$$\hat{\Sigma} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{x})'$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i$$

is the vector of sample means.

A Bit of R

```r
a <- c(3, 0, 0)
b <- c(0, 2, 0)
a %*% b  ## inner product
```

```
##      [,1]
## [1,]    0
```

Dependence and Independence

- **Linear combinations**: The vector $\mathbf{u}$ is a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ if

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

Dependence and Independence

- **Linear combinations**: The vector $\mathbf{u}$ is a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ if

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

- **Linear independence**: A set of vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ is linearly independent if the only solution to the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

  is $c_1 = c_2 = \cdots = c_k = 0$. If another solution exists, the set of vectors is linearly dependent.

Linear Dependence

- A set S of vectors is linearly dependent iff at least one of the vectors in S can be written as a linear combination of the other vectors in S.
- Linear independence is only defined for sets of vectors with the same number of elements
- Any linearly independent set of vectors in n-space contains at most n vectors.

Example 1

Are the following sets of vectors linearly independent?

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Example 1

Are the following sets of vectors linearly independent?

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Yes. $(c_3 = 0) \Rightarrow (c_2 = 0) \Rightarrow (c_1 = 0)$

Example 2

Are the following sets of vectors linearly independent?

$$\mathbf{v}_1 = \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} -2 \\ 2 \\ 4 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$$

Example 2

Are the following sets of vectors linearly independent?

$$\mathbf{v}_1 = \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} -2 \\ 2 \\ 4 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$$

No. $\mathbf{c} = \begin{pmatrix} 2 & 1 & -2 \end{pmatrix} \qquad \text{(e.g.)}$

Matrix Algebra

Matrices

- **Matrix**: A matrix is an array of mn real numbers arranged in m rows by n columns.

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

Matrices

- Vectors are special cases of matrices
  - Col vector of length k is a $k \times 1$ matrix
  - Row vector of length k is a $1 \times k$ matrix

Matrices

- Vectors are special cases of matrices
  - Col vector of length k is a $k \times 1$ matrix
  - Row vector of length k is a $1 \times k$ matrix
- Think of larger matrices as made up of row or column vectors. E.g.,

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix}$$

Matrices

- Vectors are special cases of matrices
  - Col vector of length k is a $k \times 1$ matrix
  - Row vector of length k is a $1 \times k$ matrix
- Think of larger matrices as made up of row or column vectors. E.g.,

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \cdots \\ \mathbf{b}_n \end{pmatrix}$$

Special Matrices

- Identity: $\mathbf{I_n} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$
- $\mathbf{J_n} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}$
- Zero: $\mathbf{0} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$

Special Matrices

- Diagonal: $\begin{pmatrix} a_{11} & 0 & 0 & 0 \\ 0 & a_{22} & 0 & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$

- Lower Triangular: $\begin{pmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$

- Upper Triangular: $\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$

- Triangular: Either upper triangular or lower triangular

Matrix Equality

- Let $\mathbf{A}$ and $\mathbf{B}$ be two $m \times n$ matrices. Then

$$\mathbf{A} = \mathbf{B}$$

iff

$$a_{ij} = b_{ij}$$

$\forall i \in \{1, \dots, m\}, \quad j \in \{1, \dots, n\}$

Matrix Addition

- Let $\mathbf{A}$ and $\mathbf{B}$ be two $m \times n$ matrices. Then

$$\mathbf{A}+\mathbf{B} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix}$$

- $\mathbf{A}$ and $\mathbf{B}$ must be same size – *conformable* for addition

Matrix Addition Example

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix}$$

$$\mathbf{A} + \mathbf{B} =$$

Matrix Addition Example

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix}$$

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} 2 & 4 & 4 \\ 6 & 6 & 8 \end{pmatrix}$$

```
A + B
```

```
##      [,1] [,2] [,3]
## [1,]    2    4    4
## [2,]    6    6    8
```

Scalar Multiplication

Scalar Multiplication: Given scalar c, the scalar multiplication $c\mathbf{A}$ is

$$c\mathbf{A} = c\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{pmatrix}$$

Scalar Multiplication Example

$$c = 2 \qquad \mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

$$c\mathbf{A} =$$

Scalar Multiplication Example

$$c = 2 \qquad \mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

$$c\mathbf{A} = \begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{pmatrix}$$

```r
c <- 2
c * A
```

```
##      [,1] [,2] [,3]
## [1,]    2    4    6
## [2,]    8   10   12
```

Matrix Multiplication

- **Matrix Multiplication**: If $\mathbf{A}$ is $m \times k$ and $\mathbf{B}$ is $k \times n$, then their product $\mathbf{C} = \mathbf{AB}$ is $m \times n$ matrix where

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj}$$

Matrix Multiplication

- **Matrix Multiplication**: If $\mathbf{A}$ is $m \times k$ and $\mathbf{B}$ is $k \times n$, then their product $\mathbf{C} = \mathbf{AB}$ is $m \times n$ matrix where

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj}$$

- Consider $\mathbf{A}$ to be composed of stacked rows $\mathbf{a}_i = \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix}$,

  $\mathbf{B}$ to be composed of stacked columns $\mathbf{b}_j = \begin{pmatrix} b_{1j} \\ b_{2j} \\ \dots \\ b_{mj} \end{pmatrix}$.

Then, $\mathbf{AB} = \mathbf{C}$, where

$$c_{ij} = \mathbf{a}_i \cdot \mathbf{b}_j$$

Notes on Matrix Multiplication

- To be *conformable* for multiplication, number of cols of first matrix must equal number of rows of second matrix

Notes on Matrix Multiplication

- To be *conformable* for multiplication, number of cols of first matrix must equal number of rows of second matrix
- The sizes of the matrices and product must be

$$(m \times k)(k \times n) = (m \times n)$$

Notes on Matrix Multiplication

- To be *conformable* for multiplication, number of cols of first matrix must equal number of rows of second matrix
- The sizes of the matrices and product must be

$$(m \times k)(k \times n) = (m \times n)$$

- Given $\mathbf{AB}$, say $\mathbf{B}$ is pre-multiplied by $\mathbf{A}$ or $\mathbf{B}$ is left-multiplied by $\mathbf{A}$ or $\mathbf{A}$ is post-multiplied by $\mathbf{B}$ or $\mathbf{A}$ is right-multiplied by $\mathbf{B}$

Warning!

- Commutative law for multiplication does not hold – order of multiplication matters:

$$\mathbf{AB} \neq \mathbf{BA}$$

- **AB** may exist, while **BA** does not.

Example

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

$\mathbf{AB} =$

Example

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

$$\mathbf{AB} = \begin{pmatrix} 2 & 3 \\ -2 & 2 \end{pmatrix}$$

Example

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

$$\mathbf{AB} = \begin{pmatrix} 2 & 3 \\ -2 & 2 \end{pmatrix}$$

$$\mathbf{BA} =$$

Example

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

$$\mathbf{AB} = \begin{pmatrix} 2 & 3 \\ -2 & 2 \end{pmatrix}$$

$$\mathbf{BA} = \begin{pmatrix} 1 & 7 \\ -1 & 3 \end{pmatrix}$$

Transpose

- **Transpose**: The transpose of the $m \times n$ matrix $\mathbf{A}$ is $n \times m$ matrix $\mathbf{A}^T$ (or $\mathbf{A}'$) obtained by interchanging rows and columns of $\mathbf{A}$.

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & 3 \\ 0 & 5 & -1 \end{pmatrix} \qquad \mathbf{A}^T =$$

Transpose

- **Transpose**: The transpose of the $m \times n$ matrix $\mathbf{A}$ is $n \times m$ matrix $\mathbf{A}^T$ (or $\mathbf{A}'$) obtained by interchanging rows and columns of $\mathbf{A}$.

- Examples:

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & 3 \\ 0 & 5 & -1 \end{pmatrix} \qquad \mathbf{A}^T =$$

Transpose

- **Transpose**: The transpose of the $m \times n$ matrix $\mathbf{A}$ is $n \times m$ matrix $\mathbf{A}^T$ (or $\mathbf{A}'$) obtained by interchanging rows and columns of $\mathbf{A}$.

- Examples:

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & 3 \\ 0 & 5 & -1 \end{pmatrix} \qquad \mathbf{A}^T =$$

Transpose

- **Transpose**: The transpose of the $m \times n$ matrix $\mathbf{A}$ is $n \times m$ matrix $\mathbf{A}^T$ (or $\mathbf{A}'$) obtained by interchanging rows and columns of $\mathbf{A}$.
- Examples:

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & 3 \\ 0 & 5 & -1 \end{pmatrix} \qquad \mathbf{A}^T = \begin{pmatrix} 4 & 0 \\ -2 & 5 \\ 3 & -1 \end{pmatrix}$$

Transpose

- **Transpose**: The transpose of the $m \times n$ matrix $\mathbf{A}$ is $n \times m$ matrix $\mathbf{A}^T$ (or $\mathbf{A}'$) obtained by interchanging rows and columns of $\mathbf{A}$.

- Examples:

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & 3 \\ 0 & 5 & -1 \end{pmatrix} \qquad \mathbf{A}^T = \begin{pmatrix} 4 & 0 \\ -2 & 5 \\ 3 & -1 \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} \qquad \mathbf{B}^T =$$

Transpose

- **Transpose**: The transpose of the $m \times n$ matrix $\mathbf{A}$ is $n \times m$ matrix $\mathbf{A}^T$ (or $\mathbf{A}'$) obtained by interchanging rows and columns of $\mathbf{A}$.

- Examples:

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & 3 \\ 0 & 5 & -1 \end{pmatrix} \qquad \mathbf{A}^T = \begin{pmatrix} 4 & 0 \\ -2 & 5 \\ 3 & -1 \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} \qquad \mathbf{B}^T = \begin{pmatrix} 2 & -1 & 3 \end{pmatrix}$$

Example Property

Example of $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$:

Example Property

Example of $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$:

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

$$(\mathbf{AB})^T =$$

Example Property

Example of $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$:

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

$$(\mathbf{AB})^T =$$

$$\left[\begin{pmatrix} 1 & 3 & 2 \\ 2 & -1 & 3 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 2 & 2 \\ 3 & -1 \end{pmatrix}\right]^T = \begin{pmatrix} 12 & 7 \\ 5 & -3 \end{pmatrix}$$

Example Continued

$$\mathbf{B}^T\mathbf{A}^T =$$

Example Continued

$$\mathbf{B}^T\mathbf{A}^T =$$

$$\begin{pmatrix} 0 & 2 & 3 \\ 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & -1 \\ 2 & 3 \end{pmatrix} =$$

Example Continued

$$\mathbf{B}^T\mathbf{A}^T =$$

$$\begin{pmatrix} 0 & 2 & 3 \\ 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & -1 \\ 2 & 3 \end{pmatrix} =$$

$$\begin{pmatrix} 12 & 7 \\ 5 & -3 \end{pmatrix}$$

Identity Matrix

The $n \times n$ identity matrix $\mathbf{I}_n$ has diagonal elements $= 1$ and off-diagonal elements $= 0$.

Examples:

$$\mathbf{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Systems of Linear Equations

Linear Equations

- **Linear Equation**: $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$
  a_i are parameters or coefficients. x_i are variables or unknowns.

Linear Equations

- **Linear Equation**: $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$
  a_i are parameters or coefficients. x_i are variables or unknowns.
- "Linear": only one variable per term and degree at most 1.

Linear Equations

- **Linear Equation**: $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$
  a_i are parameters or coefficients. x_i are variables or unknowns.
- "Linear": only one variable per term and degree at most 1.
  1. $\mathbb{R}^2$: line $\qquad x_2 = \frac{b}{a_2} - \frac{a_1}{a_2}x_1$

Linear Equations

- **Linear Equation**: $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$
  a_i are parameters or coefficients. x_i are variables or unknowns.
- "Linear": only one variable per term and degree at most 1.
  1. $\mathbb{R}^2$: line $\qquad x_2 = \frac{b}{a_2} - \frac{a_1}{a_2}x_1$
  2. $\mathbb{R}^3$: plane $\qquad x_3 = \frac{b}{a_3} - \frac{a_1}{a_3}x_1 - \frac{a_2}{a_3}x_2$

Linear Equations

- **Linear Equation**: $a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b$
  a_i are parameters or coefficients. x_i are variables or unknowns.
- "Linear": only one variable per term and degree at most 1.
  1. $\mathbb{R}^2$: line $\qquad x_2 = \frac{b}{a_2} - \frac{a_1}{a_2} x_1$
  2. $\mathbb{R}^3$: plane $\qquad x_3 = \frac{b}{a_3} - \frac{a_1}{a_3} x_1 - \frac{a_2}{a_3} x_2$
  3. $\mathbb{R}^n$: hyperplane

Systems of Linear Equations

Often interested in solving linear systems like

More Systems of Linear Equations

- More generally, we might have a system of m equations in n unknowns

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\
\vdots & & & & \vdots & & & \vdots & \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m
\end{array}$$

More Systems of Linear Equations

- More generally, we might have a system of m equations in n unknowns

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\
\vdots & & & & \vdots & & & & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m
\end{array}$$

- This scalar system is equivalent to the matrix equation

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

More Systems of Linear Equations

- More generally, we might have a system of m equations in n unknowns

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & \vdots & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

- This scalar system is equivalent to the matrix equation

$$\mathbf{Ax} = \mathbf{b}$$

- A **solution** to a linear system of m equations in n unknowns is a set of n numbers $x_1, x_2, \cdots, x_n$ that satisfy each of the m equations.

More Systems of Linear Equations

- More generally, we might have a system of m equations in n unknowns

$$
\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\
\vdots & & & & \vdots & & & & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m
\end{array}
$$

- This scalar system is equivalent to the matrix equation

$$\mathbf{Ax} = \mathbf{b}$$

- A **solution** to a linear system of m equations in n unknowns is a set of n numbers $x_1, x_2, \cdots, x_n$ that satisfy each of the m equations.
  1. $\mathbb{R}^2$: intersection of the lines.

More Systems of Linear Equations

- More generally, we might have a system of m equations in n unknowns

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & \vdots & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

- This scalar system is equivalent to the matrix equation

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

- A **solution** to a linear system of m equations in n unknowns is a set of n numbers $x_1, x_2, \cdots, x_n$ that satisfy each of the m equations.
    1. $\mathbb{R}^2$: intersection of the lines.
    2. $\mathbb{R}^3$: intersection of the planes.

More Systems of Linear Equations

- More generally, we might have a system of m equations in n unknowns

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & \vdots & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

- This scalar system is equivalent to the matrix equation

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

- A **solution** to a linear system of m equations in n unknowns is a set of n numbers $x_1, x_2, \cdots, x_n$ that satisfy each of the m equations.
    1. $\mathbb{R}^2$: intersection of the lines.
    2. $\mathbb{R}^3$: intersection of the planes.
    3. $\mathbb{R}^n$: intersection of the hyperplanes.

How many Solutions?

- Does a linear system have one, no, or multiple solutions?
  For a system of 2 equations in 2 unknowns (i.e., two lines):

How many Solutions?

- Does a linear system have one, no, or multiple solutions?
  For a system of 2 equations in 2 unknowns (i.e., two lines):
  1. One solution: lines intersect at exactly one point.

How many Solutions?

- Does a linear system have one, no, or multiple solutions?
  For a system of 2 equations in 2 unknowns (i.e., two lines):
  1. One solution: lines intersect at exactly one point.
  2. No solution: lines are parallel.

How many Solutions?

- Does a linear system have one, no, or multiple solutions?
  For a system of 2 equations in 2 unknowns (i.e., two lines):
  1. One solution: lines intersect at exactly one point.
  2. No solution: lines are parallel.
  3. Infinite solutions: lines coincide.

Matrices to Represent Linear Systems

Matrices are an efficient way to represent linear systems such as

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\
\vdots & & & & \vdots & & & & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m
\end{array}$$

Matrices to Represent Linear Systems

Matrices are an efficient way to represent linear systems such as

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\
\vdots & & & & \vdots & & & & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m
\end{array}$$

as $\mathbf{Ax} = \mathbf{b}$

Coefficient Matrix

The $m \times n$ **coefficient matrix A** is an array of mn real numbers arranged in m rows by n columns:

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

Variable & Output Vectors

- The unknown quantities are represented by the vector

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Variable & Output Vectors

- The unknown quantities are represented by the vector
$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$
- The RHS of the linear system is represented by the vector
$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Inverse of a Matrix

- An $n \times n$ matrix $\mathbf{A}$ is *nonsingular* or *invertible* if there exists an $n \times n$ matrix $\mathbf{A}^{-1}$ such that

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

Inverse of a Matrix

- An $n \times n$ matrix $\mathbf{A}$ is *nonsingular* or *invertible* if there exists an $n \times n$ matrix $\mathbf{A}^{-1}$ such that

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

- Then, $\mathbf{A}^{-1}$ is the inverse of $\mathbf{A}$.

Inverse of a Matrix

- An $n \times n$ matrix $\mathbf{A}$ is *nonsingular* or *invertible* if there exists an $n \times n$ matrix $\mathbf{A}^{-1}$ such that

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

- Then, $\mathbf{A}^{-1}$ is the inverse of $\mathbf{A}$.
- If there is no such $\mathbf{A}^{-1}$, then $\mathbf{A}$ is *singular* or *noninvertible*.

Example of Inverses

- Example: Let

$$\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 2 & 2 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} -1 & \frac{3}{2} \\ 1 & -1 \end{pmatrix}$$

Example of Inverses

- Example: Let

$$\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 2 & 2 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} -1 & \frac{3}{2} \\ 1 & -1 \end{pmatrix}$$

Since

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n$$

we conclude that $\mathbf{B}$ is the inverse of $\mathbf{A}$, $\mathbf{A}^{-1}$, and that $\mathbf{A}$ is nonsingular.

Calculating the Inverse in R

```r
A <- matrix(c(1, 4, 5, 6), 2, 2, byrow = TRUE)
A
```

```
##      [,1] [,2]
## [1,]    1    4
## [2,]    5    6
```

Calculating the Inverse in R

```r
A <- matrix(c(1, 4, 5, 6), 2, 2, byrow = TRUE)
A
```

```
##      [,1] [,2]
## [1,]    1    4
## [2,]    5    6
```

```r
solve(A)
```

```
##            [,1]        [,2]
## [1,] -0.4285714  0.28571429
## [2,]  0.3571429 -0.07142857
```

Calculating the Inverse in R

```r
solve(A) %*% A
```

```
##                [,1]         [,2]
## [1,]  1.000000e+00 1.110223e-16
## [2,] -4.163336e-17 1.000000e+00
```

Calculating the Inverse in R

```r
solve(A) %*% A
```

```
##                   [,1]         [,2]
## [1,]  1.000000e+00 1.110223e-16
## [2,] -4.163336e-17 1.000000e+00
```

```r
A %*% solve(A)
```

```
##                   [,1] [,2]
## [1,]  1.000000e+00    0
## [2,] -3.330669e-16    1
```

Linear Systems and Inverses

- Matrix representation of a linear system

$$\mathbf{Ax} = \mathbf{b}$$

- If $\mathbf{A}$ is an $n \times n$ matrix,then $\mathbf{Ax} = \mathbf{b}$ is a system of n equations in n unknowns.

Linear Systems and Inverses

- Matrix representation of a linear system

$$\mathbf{Ax} = \mathbf{b}$$

- If $\mathbf{A}$ is an $n \times n$ matrix,then $\mathbf{Ax} = \mathbf{b}$ is a system of n equations in n unknowns.
- Suppose $\mathbf{A}$ is nonsingular $\quad \Longrightarrow \quad \mathbf{A}^{-1}$ exists.

Linear Systems and Inverses

- Matrix representation of a linear system

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

- If $\mathbf{A}$ is an $n \times n$ matrix,then $\mathbf{A}\mathbf{x} = \mathbf{b}$ is a system of n equations in n unknowns.
- Suppose $\mathbf{A}$ is nonsingular $\quad \Longrightarrow \quad \mathbf{A}^{-1}$ exists.
- To solve this system, we can premultiply each side by $\mathbf{A}^{-1}$ and simplify:

$$\mathbf{A}^{-1}(\mathbf{A}\mathbf{x}) \;\;=\;\; \mathbf{A}^{-1}\mathbf{b}$$

Linear Systems and Inverses

- Matrix representation of a linear system

$$\mathbf{Ax} = \mathbf{b}$$

- If $\mathbf{A}$ is an $n \times n$ matrix,then $\mathbf{Ax} = \mathbf{b}$ is a system of n equations in n unknowns.
- Suppose $\mathbf{A}$ is nonsingular $\quad \Longrightarrow \quad \mathbf{A}^{-1}$ exists.
- To solve this system, we can premultiply each side by $\mathbf{A}^{-1}$ and simplify:

$$\begin{aligned} \mathbf{A}^{-1}(\mathbf{Ax}) &= \mathbf{A}^{-1}\mathbf{b} \\ (\mathbf{A}^{-1}\mathbf{A})\mathbf{x} &= \mathbf{A}^{-1}\mathbf{b} \end{aligned}$$

Linear Systems and Inverses

- Matrix representation of a linear system

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

- If $\mathbf{A}$ is an $n \times n$ matrix,then $\mathbf{A}\mathbf{x} = \mathbf{b}$ is a system of n equations in n unknowns.
- Suppose $\mathbf{A}$ is nonsingular $\quad \Longrightarrow \quad \mathbf{A}^{-1}$ exists.
- To solve this system, we can premultiply each side by $\mathbf{A}^{-1}$ and simplify:

$$\begin{aligned}
\mathbf{A}^{-1}(\mathbf{A}\mathbf{x}) &= \mathbf{A}^{-1}\mathbf{b} \\
(\mathbf{A}^{-1}\mathbf{A})\mathbf{x} &= \mathbf{A}^{-1}\mathbf{b} \\
\mathbf{I}_n\mathbf{x} &= \mathbf{A}^{-1}\mathbf{b}
\end{aligned}$$

Linear Systems and Inverses

- Matrix representation of a linear system

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

- If $\mathbf{A}$ is an $n \times n$ matrix,then $\mathbf{A}\mathbf{x} = \mathbf{b}$ is a system of n equations in n unknowns.
- Suppose $\mathbf{A}$ is nonsingular $\implies$ $\mathbf{A}^{-1}$ exists.
- To solve this system, we can premultiply each side by $\mathbf{A}^{-1}$ and simplify:

$$\begin{aligned} \mathbf{A}^{-1}(\mathbf{A}\mathbf{x}) &= \mathbf{A}^{-1}\mathbf{b} \\ (\mathbf{A}^{-1}\mathbf{A})\mathbf{x} &= \mathbf{A}^{-1}\mathbf{b} \\ \mathbf{I}_n\mathbf{x} &= \mathbf{A}^{-1}\mathbf{b} \\ \mathbf{x} &= \mathbf{A}^{-1}\mathbf{b} \end{aligned}$$

The Linear Regression Parameters

- We have linear systems with n obs and k predictors

The Linear Regression Parameters

- We have linear systems with n obs and k predictors
  - $\mathbf{X}$ is $n \times k$ matrix of predictor variable data

The Linear Regression Parameters

- We have linear systems with n obs and k predictors
  - $\mathbf{X}$ is $n \times k$ matrix of predictor variable data
  - $\mathbf{y}$ is $n \times 1$ vector of outcome variable values

The Linear Regression Parameters

- We have linear systems with n obs and k predictors
  - $\mathbf{X}$ is $n \times k$ matrix of predictor variable data
  - $\mathbf{y}$ is $n \times 1$ vector of outcome variable values
  - $\mathbf{b}$ is $k \times 1$ vector of linear parameters (β's)

The Linear Regression Parameters

- We have linear systems with n obs and k predictors
  - $\mathbf{X}$ is $n \times k$ matrix of predictor variable data
  - $\mathbf{y}$ is $n \times 1$ vector of outcome variable values
  - $\mathbf{b}$ is $k \times 1$ vector of linear parameters (β's)
- Then

$$\mathbf{Xb} = \mathbf{y}$$

The Linear Regression Parameters

- We have linear systems with n obs and k predictors
  - $\mathbf{X}$ is $n \times k$ matrix of predictor variable data
  - $\mathbf{y}$ is $n \times 1$ vector of outcome variable values
  - $\mathbf{b}$ is $k \times 1$ vector of linear parameters (β's)
- Then

$$\mathbf{Xb} = \mathbf{y}$$

- Since usually $n \gg k$, $\mathbf{X}$ not square

The Linear Regression Parameters

- We have linear systems with n obs and k predictors
  - $\mathbf{X}$ is $n \times k$ matrix of predictor variable data
  - $\mathbf{y}$ is $n \times 1$ vector of outcome variable values
  - $\mathbf{b}$ is $k \times 1$ vector of linear parameters (β's)
- Then

$$\mathbf{Xb} = \mathbf{y}$$

- Since usually $n \gg k$, $\mathbf{X}$ not square
- To isolate $\mathbf{b}$, how to make premultiplying matrix square?

The Linear Regression Parameters

- We have linear systems with n obs and k predictors
  - $\mathbf{X}$ is $n \times k$ matrix of predictor variable data
  - $\mathbf{y}$ is $n \times 1$ vector of outcome variable values
  - $\mathbf{b}$ is $k \times 1$ vector of linear parameters (β's)
- Then

$$\mathbf{X}\mathbf{b} = \mathbf{y}$$

- Since usually $n \gg k$, $\mathbf{X}$ not square
- To isolate $\mathbf{b}$, how to make premultiplying matrix square?
- Observe:

$$\underset{(n\times k)}{\mathbf{X}}\underset{(k\times n)}{\mathbf{X}'} = \underset{(n\times n)}{}$$

can't premultiply $\mathbf{b}$.

The Linear Regression Parameters

- We have linear systems with n obs and k predictors
  - $\mathbf{X}$ is $n \times k$ matrix of predictor variable data
  - $\mathbf{y}$ is $n \times 1$ vector of outcome variable values
  - $\mathbf{b}$ is $k \times 1$ vector of linear parameters (β's)
- Then

$$\mathbf{Xb} = \mathbf{y}$$

- Since usually $n \gg k$, $\mathbf{X}$ not square
- To isolate $\mathbf{b}$, how to make premultiplying matrix square?
- Observe:

$$\underset{(n\times k)}{\mathbf{X}}\underset{(k\times n)}{\mathbf{X}'} = \underset{(n\times n)}{}$$

can't premultiply $\mathbf{b}$.

The Linear Regression Parameters

- We have linear systems with n obs and k predictors
  - $\mathbf{X}$ is $n \times k$ matrix of predictor variable data
  - $\mathbf{y}$ is $n \times 1$ vector of outcome variable values
  - $\mathbf{b}$ is $k \times 1$ vector of linear parameters (β's)
- Then

$$\mathbf{Xb} = \mathbf{y}$$

- Since usually $n \gg k$, $\mathbf{X}$ not square
- To isolate $\mathbf{b}$, how to make premultiplying matrix square?
- Observe:

$$\underset{(n\times k)}{\mathbf{X}}\underset{(k\times n)}{\mathbf{X}'} = \underset{(n\times n)}{}$$

can't premultiply $\mathbf{b}$. But this can:

$$\underset{(k\times n)}{\mathbf{X}'}\underset{(n\times k)}{\mathbf{X}} = \underset{(k\times k)}{}$$

The Linear Regression Parameters

$$\mathbf{Xb} = \mathbf{y}$$

The Linear Regression Parameters

$$
\begin{aligned}
\mathbf{Xb} &= \mathbf{y} \\
\mathbf{X}'\mathbf{Xb} &= \mathbf{X}'\mathbf{y}
\end{aligned}
$$

The Linear Regression Parameters

$$\begin{aligned} \mathbf{X}\mathbf{b} &= \mathbf{y} \\ \mathbf{X}'\mathbf{X}\mathbf{b} &= \mathbf{X}'\mathbf{y} \end{aligned}$$

How to isolate $\mathbf{b}$?

The Linear Regression Parameters

$$
\begin{aligned}
\mathbf{Xb} &= \mathbf{y} \\
\mathbf{X'Xb} &= \mathbf{X'y}
\end{aligned}
$$

How to isolate $\mathbf{b}$? Multiply by an inverse.

The Linear Regression Parameters

$$\begin{aligned} \mathbf{Xb} &= \mathbf{y} \\ \mathbf{X'Xb} &= \mathbf{X'y} \end{aligned}$$

How to isolate $\mathbf{b}$? Multiply by an inverse.

$$(\mathbf{X'X})^{-1}\mathbf{X'Xb} = (\mathbf{X'X})^{\mathbf{-1}}\mathbf{X'y}$$

The Linear Regression Parameters

$$\begin{aligned} \mathbf{Xb} &= \mathbf{y} \\ \mathbf{X'Xb} &= \mathbf{X'y} \end{aligned}$$

How to isolate $\mathbf{b}$? Multiply by an inverse.

$$\begin{aligned} (\mathbf{X'X})^{-1}\mathbf{X'Xb} &= (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'y} \\ \mathbf{I}_k\mathbf{b} &= (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'y} \end{aligned}$$

The Linear Regression Parameters

$$\begin{aligned} \mathbf{Xb} &= \mathbf{y} \\ \mathbf{X'Xb} &= \mathbf{X'y} \end{aligned}$$

How to isolate $\mathbf{b}$? Multiply by an inverse.

$$\begin{aligned} (\mathbf{X'X})^{-1}\mathbf{X'Xb} &= (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'y} \\ \mathbf{I}_k\mathbf{b} &= (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'y} \\ \mathbf{b} &= (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'y} \end{aligned}$$

The Linear Regression Parameters

$$\begin{aligned}\mathbf{Xb} &= \mathbf{y} \\ \mathbf{X'Xb} &= \mathbf{X'y}\end{aligned}$$

How to isolate $\mathbf{b}$? Multiply by an inverse.

$$\begin{aligned}(\mathbf{X'X})^{-1}\mathbf{X'Xb} &= (\mathbf{X'X})^{\mathbf{-1}}\mathbf{X'y} \\ \mathbf{I}_k\mathbf{b} &= (\mathbf{X'X})^{\mathbf{-1}}\mathbf{X'y} \\ \mathbf{b} &= (\mathbf{X'X})^{\mathbf{-1}}\mathbf{X'y}\end{aligned}$$

Befriend $(\mathbf{X'X})^{\mathbf{-1}}\mathbf{X'y}$.

The Linear Regression Parameters

$$\begin{aligned}\mathbf{Xb} &= \mathbf{y}\\ \mathbf{X}'\mathbf{Xb} &= \mathbf{X}'\mathbf{y}\end{aligned}$$

How to isolate $\mathbf{b}$? Multiply by an inverse.

$$\begin{aligned}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Xb} &= (\mathbf{X}'\mathbf{X})^{\mathbf{-1}}\mathbf{X}'\mathbf{y}\\ \mathbf{I}_k\mathbf{b} &= (\mathbf{X}'\mathbf{X})^{\mathbf{-1}}\mathbf{X}'\mathbf{y}\\ \mathbf{b} &= (\mathbf{X}'\mathbf{X})^{\mathbf{-1}}\mathbf{X}'\mathbf{y}\end{aligned}$$

Befriend $(\mathbf{X}'\mathbf{X})^{\mathbf{-1}}\mathbf{X}'\mathbf{y}$. If you understand it, its cousins, and their properties (both strengths and weaknesses), your data-analytic future will be bright.

Solving a System in R ($n \times n$)

$$\mathbf{y} = \mathbf{X}\mathbf{b}$$

```r
x1 <- c(1, 3, 5)
x2 <- c(3, 1, 2)
x3 <- c(1, 1, 1)
y <-  4 * x1 + 3 * x2 + x3
(X <- cbind(x1, x2, x3))
```

```
##      x1 x2 x3
## [1,]  1  3  1
## [2,]  3  1  1
## [3,]  5  2  1
```

Solving a System in R ($n \times n$)

```
Xinv <- solve(X)
b <- Xinv %*% y
b
```

```
##    [,1]
## x1    4
## x2    3
## x3    1
```