Mathematical Preliminaries

K. Subramani[1]

[1] Lane Department of Computer Science and Electrical Engineering
West Virginia University

January 20, 2015

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization
6. Financial Mathematics
   - Quantitative models
   - Problem Types

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization
6. Financial Mathematics
   - Quantitative models
   - Problem Types

Basics

Basics

Definition

Basics

Definition

A vector is an ordered array of numbers.

Basics

Definition

A vector is an ordered array of numbers.

Geometric Representation

Basics

Definition

A vector is an ordered array of numbers.

Geometric Representation

The collection of all m-dimensional vectors is called **Euclidean *m*-space** and is denoted by E^m

Basics

Definition

A vector is an ordered array of numbers.

Geometric Representation

The collection of all m-dimensional vectors is called **Euclidean m-space** and is denoted by E^m (also $\Re^m$).

Basics

Definition

A vector is an ordered array of numbers.

Geometric Representation

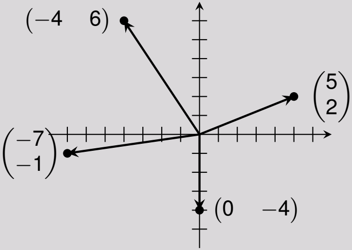
The collection of all m-dimensional vectors is called **Euclidean m-space** and is denoted by E^m (also $\Re^m$).
Vectors can be represented geometrically, where a vector can be thought of as either a point or as an arrow directed from the origin to the point.

Example

Example

Example

Euclidean 2-space, E^2.



Vector Addition

Vector Addition

Vectors of the same type (row or column) can be added if they have the same number of entries.

Vector Addition

Vector Addition

Vectors of the same type (row or column) can be added if they have the same number of entries.
Given two vectors **a** and **b**, we simply add one element in **a** with the corresponding element in **b** that is in the same position.

Vector Addition

Vector Addition

Vectors of the same type (row or column) can be added if they have the same number of entries.
Given two vectors **a** and **b**, we simply add one element in **a** with the corresponding element in **b** that is in the same position.
In other words, given $\mathbf{c} = \mathbf{a} + \mathbf{b}$ where c_i is the element in the ith position, we have $c_i = a_i + b_i$.

Vector Addition

Vector Addition

Vectors of the same type (row or column) can be added if they have the same number of entries.
Given two vectors **a** and **b**, we simply add one element in **a** with the corresponding element in **b** that is in the same position.
In other words, given $\mathbf{c} = \mathbf{a} + \mathbf{b}$ where c_i is the element in the ith position, we have $c_i = a_i + b_i$.

Vector addition satisfies both the commutative ($\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$) and associative ($\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + \mathbf{b} + \mathbf{c}$) laws.

Vector Addition Example

$$\mathbf{a} = \begin{pmatrix} 4 \\ 0 \\ 7 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 5 \\ 9 \\ 1 \end{pmatrix} \quad \mathbf{c} = \begin{pmatrix} 6 & 8 & 0 \end{pmatrix} \quad \mathbf{d} = \begin{pmatrix} 4 \\ 10 \\ 2 \\ 3 \end{pmatrix}$$

Vector Addition Example

$$\mathbf{a} = \begin{pmatrix} 4 \\ 0 \\ 7 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 5 \\ 9 \\ 1 \end{pmatrix} \quad \mathbf{c} = \begin{pmatrix} 6 & 8 & 0 \end{pmatrix} \quad \mathbf{d} = \begin{pmatrix} 4 \\ 10 \\ 2 \\ 3 \end{pmatrix}$$

$$\mathbf{a} + \mathbf{b} = \begin{pmatrix} 4 \\ 0 \\ 7 \end{pmatrix} + \begin{pmatrix} 5 \\ 9 \\ 1 \end{pmatrix} = \begin{pmatrix} 9 \\ 9 \\ 8 \end{pmatrix}$$

$\mathbf{a} + \mathbf{c}$ is undefined (not the same type)
$\mathbf{a} + \mathbf{d}$ is undefined (different number of elements)

Scalar Multiplication

Multiplication of a Vector by a Scalar

We define a **scalar** as an element of E^1, Euclidean 1-space.

Scalar Multiplication

Multiplication of a Vector by a Scalar

We define a **scalar** as an element of E^1, Euclidean 1-space. For example, $3, 19, 37.5$, and $\frac{2}{3}$ are scalars.

Scalar Multiplication

Multiplication of a Vector by a Scalar

We define a **scalar** as an element of E^1, Euclidean 1-space. For example, $3, 19, 37.5$, and $\frac{2}{3}$ are scalars.
To multiply a vector by a scalar, we simply multiply each element in the vector by the scalar.

Scalar Multiplication

Multiplication of a Vector by a Scalar

We define a **scalar** as an element of E^1, Euclidean 1-space. For example, $3, 19, 37.5$, and $\frac{2}{3}$ are scalars.
To multiply a vector by a scalar, we simply multiply each element in the vector by the scalar.
For example, if we are given a scalar α, a row vector **a**, and a column vector **b**, we have

Scalar Multiplication

Multiplication of a Vector by a Scalar

We define a **scalar** as an element of E^1, Euclidean 1-space. For example, $3, 19, 37.5$, and $\frac{2}{3}$ are scalars.
To multiply a vector by a scalar, we simply multiply each element in the vector by the scalar.
For example, if we are given a scalar α, a row vector **a**, and a column vector **b**, we have

$$\alpha \cdot \mathbf{a} = \alpha \cdot (a_1, a_2, \ldots, a_n) = (\alpha \cdot a_1, \alpha \cdot a_2, \ldots, \alpha \cdot a_n)$$

Scalar Multiplication

Multiplication of a Vector by a Scalar

We define a **scalar** as an element of E^1, Euclidean 1-space. For example, $3, 19, 37.5$, and $\frac{2}{3}$ are scalars.
To multiply a vector by a scalar, we simply multiply each element in the vector by the scalar.
For example, if we are given a scalar α, a row vector **a**, and a column vector **b**, we have

$$\alpha \cdot \mathbf{a} = \alpha \cdot (a_1, a_2, \ldots, a_n) = (\alpha \cdot a_1, \alpha \cdot a_2, \ldots, \alpha \cdot a_n)$$

$$\alpha \cdot \mathbf{b} = \alpha \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = \begin{pmatrix} \alpha \cdot b_1 \\ \alpha \cdot b_2 \\ \vdots \\ \alpha \cdot b_m \end{pmatrix}$$

Vector Multiplication

Vector Multiplication

We can multiply two vectors if both have the same number of entries, one of them is a row vector, and the other is a column vector.

Vector Multiplication

Vector Multiplication

We can multiply two vectors if both have the same number of entries, one of them is a row vector, and the other is a column vector.
The result, often called the **dot product**, is a scalar.

Vector Multiplication

Vector Multiplication

We can multiply two vectors if both have the same number of entries, one of them is a row vector, and the other is a column vector.
The result, often called the **dot product**, is a scalar. By convention, having $\mathbf{a} \cdot \mathbf{b}$ or $\mathbf{ab}$ means $\mathbf{a}$ is the row vector and $\mathbf{b}$ is the column vector.

Vector Multiplication

Vector Multiplication

We can multiply two vectors if both have the same number of entries, one of them is a row vector, and the other is a column vector.
The result, often called the **dot product**, is a scalar. By convention, having $\mathbf{a} \cdot \mathbf{b}$ or $\mathbf{ab}$ means $\mathbf{a}$ is the row vector and $\mathbf{b}$ is the column vector.
To multiply the vectors, we multiply the corresponding entries and add the results.

Vector Multiplication

Vector Multiplication

We can multiply two vectors if both have the same number of entries, one of them is a row vector, and the other is a column vector.

The result, often called the **dot product**, is a scalar. By convention, having $\mathbf{a} \cdot \mathbf{b}$ or $\mathbf{ab}$ means $\mathbf{a}$ is the row vector and $\mathbf{b}$ is the column vector.

To multiply the vectors, we multiply the corresponding entries and add the results.

What this means that if we assume the vectors have m entries, we have

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{ab} = \sum_{i=1}^{m} a_i b_i = \alpha.$$

Vector Multiplication

Vector Multiplication

We can multiply two vectors if both have the same number of entries, one of them is a row vector, and the other is a column vector.
The result, often called the **dot product**, is a scalar. By convention, having $\mathbf{a} \cdot \mathbf{b}$ or $\mathbf{ab}$ means $\mathbf{a}$ is the row vector and $\mathbf{b}$ is the column vector.
To multiply the vectors, we multiply the corresponding entries and add the results.
What this means that if we assume the vectors have m entries, we have

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{ab} = \sum_{i=1}^{m} a_i b_i = \alpha.$$

We should also note that vector multiplication satisfies the distributive law $\mathbf{a}(\mathbf{b} + \mathbf{c}) = \mathbf{ab} + \mathbf{ac}$.

Vectors

Vector Multiplication Example

$$\mathbf{a} = \begin{pmatrix} 3 \\ 0 \\ 7 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} -2 \\ 10 \\ 1 \end{pmatrix} \quad \mathbf{c} = \begin{pmatrix} 4 & 9 & 2 \end{pmatrix}$$

$$\mathbf{d} = \begin{pmatrix} 5 & 1 & 4 & 2 \end{pmatrix} \quad \mathbf{e} = \begin{pmatrix} 3 & -2 \end{pmatrix}$$

Vectors

Vector Multiplication Example

$$\mathbf{a} = \begin{pmatrix} 3 \\ 0 \\ 7 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} -2 \\ 10 \\ 1 \end{pmatrix} \quad \mathbf{c} = \begin{pmatrix} 4 & 9 & 2 \end{pmatrix}$$

$$\mathbf{d} = \begin{pmatrix} 5 & 1 & 4 & 2 \end{pmatrix} \quad \mathbf{e} = \begin{pmatrix} 3 & -2 \end{pmatrix}$$

$$\mathbf{ca} = \begin{pmatrix} 4 & 9 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ 7 \end{pmatrix} = 12 + 0 + 14 = 26$$

$$\mathbf{cb} = \begin{pmatrix} 4 & 9 & 2 \end{pmatrix} \begin{pmatrix} -2 \\ 10 \\ 1 \end{pmatrix} = -8 + 90 + 2 = 84$$

Norms

Norm of a Vector

The $\mathbf{L_p}$ norm of a vector $\mathbf{a} \in E^n$, denoted by $\|\mathbf{a}\|_p$, is a measure of the size of $\mathbf{a}$ and is given by $\|\mathbf{a}\|_p = \left(\sum_{i=1}^{n} |a_i|^p\right)^{1/p}$.

Norms

Norm of a Vector

The $\mathbf{L_p}$ norm of a vector $\mathbf{a} \in E^n$, denoted by $\|\mathbf{a}\|_p$, is a measure of the size of $\mathbf{a}$ and is given by $\|\mathbf{a}\|_p = \left(\sum_{i=1}^{n} |a_i|^p\right)^{1/p}$.

Some common norms are the L_1 norm

Norms

Norm of a Vector

The $\mathbf{L_p}$ norm of a vector $\mathbf{a} \in E^n$, denoted by $||\mathbf{a}||_p$, is a measure of the size of $\mathbf{a}$ and is given by $||\mathbf{a}||_p = \left(\sum_{i=1}^{n} |a_i|^p\right)^{1/p}$.

Some common norms are the L_1 norm (Manhattan),

Norms

Norm of a Vector

The $\mathbf{L_p}$ norm of a vector $\mathbf{a} \in E^n$, denoted by $\|\mathbf{a}\|_p$, is a measure of the size of $\mathbf{a}$ and is given by $\|\mathbf{a}\|_p = \left(\sum_{i=1}^{n} |a_i|^p \right)^{1/p}$.

Some common norms are the L_1 norm (Manhattan), L_2 norm

Norms

Norm of a Vector

The $\mathbf{L_p}$ norm of a vector $\mathbf{a} \in E^n$, denoted by $||\mathbf{a}||_p$, is a measure of the size of $\mathbf{a}$ and is given by $||\mathbf{a}||_p = \left(\sum_{i=1}^{n} |a_i|^p\right)^{1/p}$.

Some common norms are the L_1 norm (Manhattan), L_2 norm (Euclidean)

Norms

Norm of a Vector

The $\mathbf{L_p}$ norm of a vector $\mathbf{a} \in E^n$, denoted by $||\mathbf{a}||_p$, is a measure of the size of $\mathbf{a}$ and is given by $||\mathbf{a}||_p = \left(\sum_{i=1}^{n} |a_i|^p \right)^{1/p}$.

Some common norms are the L_1 norm (Manhattan), L_2 norm (Euclidean) ant the L_∞ norm.

Norms

Norm of a Vector

The $\mathbf{L_p}$ norm of a vector $\mathbf{a} \in E^n$, denoted by $\|\mathbf{a}\|_p$, is a measure of the size of $\mathbf{a}$ and is given by $\|\mathbf{a}\|_p = \left(\sum_{i=1}^{n} |a_i|^p\right)^{1/p}$.

Some common norms are the L_1 norm (Manhattan), L_2 norm (Euclidean) ant the L_∞ norm.

Example

$$\mathbf{a} = \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} \quad \|\mathbf{a}\|_2 = [3^2 + 2^2 + (-1)^2]^{1/2} = (14)^{1/2}$$

Norms

Norm of a Vector

The $\mathbf{L_p}$ norm of a vector $\mathbf{a} \in E^n$, denoted by $\|\mathbf{a}\|_p$, is a measure of the size of **a** and is given by $\|\mathbf{a}\|_p = \left(\sum_{i=1}^{n} |a_i|^p \right)^{1/p}$.

Some common norms are the L_1 norm (Manhattan), L_2 norm (Euclidean) ant the L_∞ norm.

Example

$$\mathbf{a} = \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} \qquad \|\mathbf{a}\|_2 = [3^2 + 2^2 + (-1)^2]^{1/2} = (14)^{1/2}$$

Note

The dot product of two vectors can also be defined by using the Euclidean norm, which is given by $\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\|_2 \cdot \|\mathbf{b}\|_2 \cos\theta$, where θ is the angle between the two vectors.

Vectors

Special Vector Types

Unit Vector - Has a 1 in the j^{th} position and 0's elsewhere.

Vectors

Special Vector Types

Unit Vector - Has a 1 in the j^{th} position and 0's elsewhere. We normally denote this by $\mathbf{e}_j$, where 1 appears in the j^{th} position.

Vectors

Special Vector Types

Unit Vector - Has a 1 in the j^{th} position and 0's elsewhere. We normally denote this by $\mathbf{e}_j$, where 1 appears in the j^{th} position.
For example, if $\mathbf{e}_j \in E^3$,

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Vectors

Special Vector Types

Unit Vector - Has a 1 in the j^{th} position and 0's elsewhere. We normally denote this by $\mathbf{e}_j$, where 1 appears in the j^{th} position.
For example, if $\mathbf{e}_j \in E^3$,

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Null or Zero Vector - Denoted by $\mathbf{0}$, is a vector having only 0's.

Vectors

Special Vector Types

Unit Vector - Has a 1 in the j^{th} position and 0's elsewhere. We normally denote this by $\mathbf{e}_j$, where 1 appears in the j^{th} position.
For example, if $\mathbf{e}_j \in E^3$,

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Null or Zero Vector - Denoted by **0**, is a vector having only 0's.

Sum Vector - Denoted by **1**, is a vector having only 1's.

Vectors

Special Vector Types

Unit Vector - Has a 1 in the j^{th} position and 0's elsewhere. We normally denote this by $\mathbf{e}_j$, where 1 appears in the j^{th} position.
For example, if $\mathbf{e}_j \in E^3$,

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Null or Zero Vector - Denoted by **0**, is a vector having only 0's.

Sum Vector - Denoted by **1**, is a vector having only 1's.
We call this the sum vector because the dot product of **1** and some vector **a** is a scalar that is equal to the sum of the elements in **a**.

Vectors

Special Vector Types

Unit Vector - Has a 1 in the j^{th} position and 0's elsewhere. We normally denote this by $\mathbf{e}_j$, where 1 appears in the j^{th} position.
For example, if $\mathbf{e}_j \in E^3$,

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Null or Zero Vector - Denoted by $\mathbf{0}$, is a vector having only 0's.

Sum Vector - Denoted by $\mathbf{1}$, is a vector having only 1's.
We call this the sum vector because the dot product of $\mathbf{1}$ and some vector $\mathbf{a}$ is a scalar that is equal to the sum of the elements in $\mathbf{a}$.

Vectors

Linear Dependence and Independence

A set of vectors, $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ is **linearly dependent** if there exist some scalars, α_i, that are not all zero such that

$$\alpha_1 \cdot \mathbf{a}_1 + \alpha_2 \cdot \mathbf{a}_2 + \cdots + \alpha_m \cdot \mathbf{a}_m = \mathbf{0} \quad (1)$$

Vectors

Linear Dependence and Independence

A set of vectors, $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ is **linearly dependent** if there exist some scalars, α_i, that are not all zero such that

$$\alpha_1 \cdot \mathbf{a}_1 + \alpha_2 \cdot \mathbf{a}_2 + \cdots + \alpha_m \cdot \mathbf{a}_m = \mathbf{0} \quad (1)$$

If the only set of scalars, α_i, for which the above equation holds is $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$, the vectors are **linearly independent**.

Example

Example

Example

Linearly Dependent:

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \mathbf{a}_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \mathbf{a}_3 = \begin{pmatrix} 8 \\ 11 \end{pmatrix}$$

$$2\mathbf{a}_1 + 3\mathbf{a}_2 - 1\mathbf{a}_3 = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 3\begin{pmatrix} 2 \\ 3 \end{pmatrix} - 1\begin{pmatrix} 8 \\ 11 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Vectors

Example

Linearly Independent:

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \mathbf{a}_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

Consider the equation

$$\alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\alpha_1 + 2\alpha_2 = 0 \tag{2}$$

$$\alpha_1 = 0 \tag{3}$$

We can see that the only solution is $\alpha_1 = \alpha_2 = 0$. This means $\mathbf{a}_1$ and $\mathbf{a}_2$ are linearly independent.

Vectors

Spanning Sets and Bases

The vectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p \in E^n$ are said to form a **spanning set** if every vector in E^n can be written as a linear combination of the $\mathbf{b}_i$.

Vectors

Spanning Sets and Bases

The vectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p \in E^n$ are said to form a **spanning set** if every vector in E^n can be written as a linear combination of the $\mathbf{b}_i$.
In other words, if $\mathbf{v} \in E^n$, then there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_p$ such that $\mathbf{v} = \alpha_1 \cdot \mathbf{b}_1 + \alpha_2 \cdot \mathbf{b}_2 + \cdots + \alpha_p \cdot \mathbf{b}_p$.

Vectors

Spanning Sets and Bases

The vectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p \in E^n$ are said to form a **spanning set** if every vector in E^n can be written as a linear combination of the $\mathbf{b}_i$.
In other words, if $\mathbf{v} \in E^n$, then there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_p$ such that
$\mathbf{v} = \alpha_1 \cdot \mathbf{b}_1 + \alpha_2 \cdot \mathbf{b}_2 + \cdots + \alpha_p \cdot \mathbf{b}_p$.
We say that the vectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n \in E^n$ form a **basis** for E^n, if they are linearly independent and form a spanning set for E^n.

Vectors

Spanning Sets and Bases

The vectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p \in E^n$ are said to form a **spanning set** if every vector in E^n can be written as a linear combination of the $\mathbf{b}_i$.
In other words, if $\mathbf{v} \in E^n$, then there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_p$ such that
$\mathbf{v} = \alpha_1 \cdot \mathbf{b}_1 + \alpha_2 \cdot \mathbf{b}_2 + \cdots + \alpha_p \cdot \mathbf{b}_p$.
We say that the vectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n \in E^n$ form a **basis** for E^n, if they are linearly independent and form a spanning set for E^n.

Note that a basis is a minimal spanning set.

Vectors

Spanning Sets and Bases

The vectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p \in E^n$ are said to form a **spanning set** if every vector in E^n can be written as a linear combination of the $\mathbf{b}_i$.
In other words, if $\mathbf{v} \in E^n$, then there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_p$ such that
$\mathbf{v} = \alpha_1 \cdot \mathbf{b}_1 + \alpha_2 \cdot \mathbf{b}_2 + \cdots + \alpha_p \cdot \mathbf{b}_p$.
We say that the vectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n \in E^n$ form a **basis** for E^n, if they are linearly independent and form a spanning set for E^n.

Note that a basis is a minimal spanning set. This is because adding a new vector would make the set linearly dependent and removing one of the vectors would mean the remaining ones no longer span E^n.

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization
6. Financial Mathematics
   - Quantitative models
   - Problem Types

Matrices

Definition

A **matrix** is a rectangular array of numbers.

Matrices

Definition

A **matrix** is a rectangular array of numbers.
We represent them by uppercase boldface type with m rows and n columns.

Matrices

Definition

A **matrix** is a rectangular array of numbers.
We represent them by uppercase boldface type with *m* rows and *n* columns.
The **order** of a matrix is the number of rows and columns of the matrix, so the example below would be an *m* x *n* matrix.

Matrices

Definition

A **matrix** is a rectangular array of numbers.
We represent them by uppercase boldface type with *m* rows and *n* columns.
The **order** of a matrix is the number of rows and columns of the matrix, so the example below would be an *m* x *n* matrix.

Example

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}$$

Matrix Addition

Matrix Addition

If two matrices are of the same order, then we can add them together.

Matrix Addition

Matrix Addition

If two matrices are of the same order, then we can add them together.
To add two matrices, we add the elements in each corresponding position.

Matrix Addition

Matrix Addition

If two matrices are of the same order, then we can add them together.
To add two matrices, we add the elements in each corresponding position.
For example, if $\mathbf{C} = \mathbf{A} + \mathbf{B}$, then $c_{i,j} = a_{i,j} + b_{i,j}$ for every i and j.

Matrix Addition

Matrix Addition

If two matrices are of the same order, then we can add them together.
To add two matrices, we add the elements in each corresponding position.
For example, if $\mathbf{C} = \mathbf{A} + \mathbf{B}$, then $c_{i,j} = a_{i,j} + b_{i,j}$ for every i and j.
Matrix addition satisfies both the commutative and associative laws.

Matrix Addition

Matrix Addition

If two matrices are of the same order, then we can add them together.
To add two matrices, we add the elements in each corresponding position.
For example, if $\mathbf{C} = \mathbf{A} + \mathbf{B}$, then $c_{i,j} = a_{i,j} + b_{i,j}$ for every i and j.
Matrix addition satisfies both the commutative and associative laws.

Example

$$\mathbf{A} = \begin{pmatrix} 7 & 1 & -2 \\ 3 & 3 & 0 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 2 & -3 & 4 \\ 1 & 5 & 9 \end{pmatrix} \quad \mathbf{C} = \begin{pmatrix} 2 & 1 \\ 7 & 3 \\ 9 & 2 \end{pmatrix}$$

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

Scalar Multiplication

Multiplication by a Scalar

Like vectors, if we have a scalar α and a matrix **A**, the product $\alpha \cdot \mathbf{A}$ is obtained by multiplying each elements $a_{i,j}$ by α.

Scalar Multiplication

Multiplication by a Scalar

Like vectors, if we have a scalar α and a matrix $\mathbf{A}$, the product $\alpha \cdot \mathbf{A}$ is obtained by multiplying each elements $a_{i,j}$ by α.

$$\alpha \cdot \mathbf{A} = \begin{pmatrix} \alpha a_{1,1} & \alpha a_{1,2} & \cdots & \alpha a_{1,n} \\ \alpha a_{2,1} & \alpha a_{2,2} & \cdots & \alpha a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m,1} & \alpha a_{m,2} & \cdots & \alpha a_{m,n} \end{pmatrix}$$

Scalar Multiplication

Multiplication by a Scalar

Like vectors, if we have a scalar α and a matrix $\mathbf{A}$, the product $\alpha \cdot \mathbf{A}$ is obtained by multiplying each elements $a_{i,j}$ by α.

$$\alpha \cdot \mathbf{A} = \begin{pmatrix} \alpha a_{1,1} & \alpha a_{1,2} & \cdots & \alpha a_{1,n} \\ \alpha a_{2,1} & \alpha a_{2,2} & \cdots & \alpha a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m,1} & \alpha a_{m,2} & \cdots & \alpha a_{m,n} \end{pmatrix}$$

Example

$$\beta = 3 \quad \mathbf{A} = \begin{pmatrix} 8 & 3 \\ -1 & 2 \\ 7 & 1 \end{pmatrix} \quad \beta \cdot \mathbf{A} = 3 \begin{pmatrix} 8 & 3 \\ -1 & 2 \\ 7 & 1 \end{pmatrix} = \begin{pmatrix} 24 & 9 \\ -3 & 6 \\ 21 & 3 \end{pmatrix}$$

Matrix multiplicatoin

Matrix Multiplication

Two matrices **A** and **B** can be multiplied if and only if the number of columns in **A** is equal to the number of rows in **B**.

Matrix multiplicatoin

Matrix Multiplication

Two matrices **A** and **B** can be multiplied if and only if the number of columns in **A** is equal to the number of rows in **B**.
If **A** is an m x n matrix, and **B** is a p x q matrix, then $\mathbf{AB} = \mathbf{C}$ is defined as an m x q matrix if and only if $n = p$.

Matrix multiplicatoin

Matrix Multiplication

Two matrices **A** and **B** can be multiplied if and only if the number of columns in **A** is equal to the number of rows in **B**.
If **A** is an m x n matrix, and **B** is a p x q matrix, then **AB** $=$ **C** is defined as an m x q matrix if and only if $n = p$.

Each element in **C** is given by $c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$, where n is the number of columns of **A** or rows of **B**, $i = 1, \ldots, m$ where m is the number of rows of **A**, and $j = 1, \ldots, q$ where q is the number of columns of **B**.

Matrix multiplicatoin

Matrix Multiplication

Two matrices **A** and **B** can be multiplied if and only if the number of columns in **A** is equal to the number of rows in **B**.
If **A** is an m x n matrix, and **B** is a p x q matrix, then $\mathbf{AB} = \mathbf{C}$ is defined as an m x q matrix if and only if $n = p$.

Each element in **C** is given by $c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$, where n is the number of columns of **A** or rows of **B**, $i = 1, \ldots, m$ where m is the number of rows of **A**, and $j = 1, \ldots, q$ where q is the number of columns of **B**.

Matrix multiplication satisfies the associative and distributive laws, but it does **not** satisfy the commutative law in general.

Example

Example

Example

$$\mathbf{A} = \begin{pmatrix} 7 & 1 \\ 4 & -3 \\ 2 & 0 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 2 & 1 & 7 \\ 0 & -1 & 4 \end{pmatrix}$$

$$\mathbf{AB} = \begin{pmatrix} 7 & 1 \\ 4 & -3 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & 7 \\ 0 & -1 & 4 \end{pmatrix} = \begin{pmatrix} 14 & 6 & 53 \\ 8 & 7 & 16 \\ 4 & 2 & 14 \end{pmatrix}$$

Special Matrices

Special Matrices

Diagonal Matrix - A square matrix ($m = n$) whose entries that are not on the diagonal are zero.

Special Matrices

Special Matrices

Diagonal Matrix - A square matrix ($m = n$) whose entries that are not on the diagonal are zero.

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & 0 & 0 \\ 0 & a_{2,2} & 0 \\ 0 & 0 & a_{3,3} \end{pmatrix}$$

Special Matrices

Special Matrices

Diagonal Matrix - A square matrix ($m = n$) whose entries that are not on the diagonal are zero.

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & 0 & 0 \\ 0 & a_{2,2} & 0 \\ 0 & 0 & a_{3,3} \end{pmatrix}$$

Identity Matrix - A diagonal matrix where all diagonal elements are equal to 1.

Special Matrices

Special Matrices

Diagonal Matrix - A square matrix ($m = n$) whose entries that are not on the diagonal are zero.

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & 0 & 0 \\ 0 & a_{2,2} & 0 \\ 0 & 0 & a_{3,3} \end{pmatrix}$$

Identity Matrix - A diagonal matrix where all diagonal elements are equal to 1. We denote this matrix as $\mathbf{I}_m$ or $\mathbf{I}$.

Special Matrices

Special Matrices

Diagonal Matrix - A square matrix ($m = n$) whose entries that are not on the diagonal are zero.

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & 0 & 0 \\ 0 & a_{2,2} & 0 \\ 0 & 0 & a_{3,3} \end{pmatrix}$$

Identity Matrix - A diagonal matrix where all diagonal elements are equal to 1. We denote this matrix as $\mathbf{I}_m$ or $\mathbf{I}$.

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

Special Matrices

Special Matrices

Diagonal Matrix - A square matrix ($m = n$) whose entries that are not on the diagonal are zero.

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & 0 & 0 \\ 0 & a_{2,2} & 0 \\ 0 & 0 & a_{3,3} \end{pmatrix}$$

Identity Matrix - A diagonal matrix where all diagonal elements are equal to 1. We denote this matrix as $\mathbf{I}_m$ or $\mathbf{I}$.

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

Null or Zero Matrix - All elements are equal to zero and is denoted as $\mathbf{0}$.

Special Matrices

Special Matrices

Diagonal Matrix - A square matrix ($m = n$) whose entries that are not on the diagonal are zero.

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & 0 & 0 \\ 0 & a_{2,2} & 0 \\ 0 & 0 & a_{3,3} \end{pmatrix}$$

Identity Matrix - A diagonal matrix where all diagonal elements are equal to 1. We denote this matrix as $\mathbf{I}_m$ or $\mathbf{I}$.

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

Null or Zero Matrix - All elements are equal to zero and is denoted as $\mathbf{0}$. Note that this does not have to be a square matrix.

Special Matrices

Special Matrices

Diagonal Matrix - A square matrix ($m = n$) whose entries that are not on the diagonal are zero.

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & 0 & 0 \\ 0 & a_{2,2} & 0 \\ 0 & 0 & a_{3,3} \end{pmatrix}$$

Identity Matrix - A diagonal matrix where all diagonal elements are equal to 1. We denote this matrix as $\mathbf{I}_m$ or $\mathbf{I}$.

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

Null or Zero Matrix - All elements are equal to zero and is denoted as $\mathbf{0}$. Note that this does not have to be a square matrix.

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Special Matrices

Special Matrices (Contd.)

Matrix Transpose - The transpose of $\mathbf{A}$, denoted as $\mathbf{A}^t$, is a reordering of $\mathbf{A}$ by interchanging the rows and columns.

Special Matrices

Special Matrices (Contd.)

Matrix Transpose - The transpose of **A**, denoted as $\mathbf{A}^t$, is a reordering of **A** by interchanging the rows and columns. For example, row 1 of **A** would be column 1 of $\mathbf{A}^t$.

Special Matrices

Special Matrices (Contd.)

Matrix Transpose - The transpose of **A**, denoted as $\mathbf{A}^t$, is a reordering of **A** by interchanging the rows and columns. For example, row 1 of **A** would be column 1 of $\mathbf{A}^t$.

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \quad \mathbf{A}^t = \begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{m,n} \end{pmatrix}$$

Special Matrices

Special Matrices (Contd.)

Matrix Transpose - The transpose of $\mathbf{A}$, denoted as $\mathbf{A}^t$, is a reordering of $\mathbf{A}$ by interchanging the rows and columns. For example, row 1 of $\mathbf{A}$ would be column 1 of $\mathbf{A}^t$.

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \quad \mathbf{A}^t = \begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{m,n} \end{pmatrix}$$

Symmetric Matrix - A matrix $\mathbf{A}$ where $\mathbf{A} = \mathbf{A}^t$.

Special Matrices

Special Matrices (Contd.)

Matrix Transpose - The transpose of $\mathbf{A}$, denoted as $\mathbf{A}^t$, is a reordering of $\mathbf{A}$ by interchanging the rows and columns. For example, row 1 of $\mathbf{A}$ would be column 1 of $\mathbf{A}^t$.

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \quad \mathbf{A}^t = \begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{m,n} \end{pmatrix}$$

Symmetric Matrix - A matrix $\mathbf{A}$ where $\mathbf{A} = \mathbf{A}^t$.

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 6 & 4 \\ 3 & 4 & 9 \end{pmatrix}$$

Positive Semidefinite - A symmetric matrix $\mathbf{A}$ is said to be positive semidefinite, if $\mathbf{x}^{\mathbf{T}} \cdot \mathbf{A} \cdot \mathbf{x} \geq 0$ for all $\mathbf{x}$ and $\mathbf{x}^{\mathbf{T}} \cdot A \cdot \mathbf{x} = 0$, only if $\mathbf{x} = \mathbf{0}$.

Matrices

Special Matrices (Contd.)

Augmented Matrix - A matrix where the rows and columns of another matrix are appended to the original matrix.

Matrices

Special Matrices (Contd.)

Augmented Matrix - A matrix where the rows and columns of another matrix are appended to the original matrix. If **A** is augmented with **B**, we get (**A**, **B**) or (**A**|**B**).

Matrices

Special Matrices (Contd.)

Augmented Matrix - A matrix where the rows and columns of another matrix are appended to the original matrix. If $\mathbf{A}$ is augmented with $\mathbf{B}$, we get $(\mathbf{A}, \mathbf{B})$ or $(\mathbf{A}|\mathbf{B})$.

$$\mathbf{A} = \begin{pmatrix} 1 & 4 \\ 5 & 6 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 3 & 2 \\ 1 & 9 \end{pmatrix} \quad (\mathbf{A}|\mathbf{B}) = \left(\begin{array}{cc|cc} 1 & 4 & 3 & 2 \\ 5 & 6 & 1 & 9 \end{array} \right)$$

Determinants

Determinants

Given a square matrix **A**, the **determinant** denoted by $|\mathbf{A}|$ is a number associated with **A**.

Determinants

Determinants

Given a square matrix **A**, the **determinant** denoted by $|\mathbf{A}|$ is a number associated with **A**.
Determinant of a 1 x 1 matrix: $|a_{1,1}| = a_{1,1}$

Determinants

Determinants

Given a square matrix **A**, the **determinant** denoted by $|\mathbf{A}|$ is a number associated with **A**.

Determinant of a 1 x 1 matrix: $|a_{1,1}| = a_{1,1}$

Determinant of a 2 x 2 matrix: $\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$

Determinants

Determinants

Given a square matrix **A**, the **determinant** denoted by $|\mathbf{A}|$ is a number associated with **A**.

Determinant of a 1 x 1 matrix: $|a_{1,1}| = a_{1,1}$

Determinant of a 2 x 2 matrix: $\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$

Every element of a determinant, except for a 1 x 1 matrix, has an associated **minor**.

Determinants

Determinants

Given a square matrix **A**, the **determinant** denoted by $|\mathbf{A}|$ is a number associated with **A**.

Determinant of a 1 x 1 matrix: $|a_{1,1}| = a_{1,1}$

Determinant of a 2 x 2 matrix: $\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$

Every element of a determinant, except for a 1 x 1 matrix, has an associated **minor**.

To get the minor, we remove the row and column corresponding to the element and find the determinant of the new matrix.

Determinants

Determinants

Given a square matrix **A**, the **determinant** denoted by $|\mathbf{A}|$ is a number associated with **A**.

Determinant of a 1 x 1 matrix: $|a_{1,1}| = a_{1,1}$

Determinant of a 2 x 2 matrix: $\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$

Every element of a determinant, except for a 1 x 1 matrix, has an associated **minor**.

To get the minor, we remove the row and column corresponding to the element and find the determinant of the new matrix.

We denote the minor of an element $a_{i,j}$ in matrix **A** as $|\mathbf{A}_{i,j}|$.

Determinants

Determinants

Given a square matrix **A**, the **determinant** denoted by $|\mathbf{A}|$ is a number associated with **A**.

Determinant of a 1 x 1 matrix: $|a_{1,1}| = a_{1,1}$

Determinant of a 2 x 2 matrix: $\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$

Every element of a determinant, except for a 1 x 1 matrix, has an associated **minor**.

To get the minor, we remove the row and column corresponding to the element and find the determinant of the new matrix.

We denote the minor of an element $a_{i,j}$ in matrix **A** as $|\mathbf{A}_{i,j}|$.

The **cofactor** of an element is its minor with the sign $(-1)^{i+j}$ attached to it.

Example

$$|\mathbf{A}| = \begin{vmatrix} 7 & -1 & 0 \\ 3 & 2 & 1 \\ 8 & 1 & -4 \end{vmatrix}$$

The cofactor for $a_{2,1} = 3$ is

$$(-1)^{2+1}|\mathbf{A}_{2,1}| = (-1)\begin{vmatrix} -1 & 0 \\ 1 & -4 \end{vmatrix} = -4$$

Value of a determinant

Value of Determinants

The value of a determinant of order *n* is found by adding the products of each element by its respective cofactor.

Value of a determinant

Value of Determinants

The value of a determinant of order n is found by adding the products of each element by its respective cofactor. For any row i, this would be

$$|\mathbf{A}| = \sum_{j=1}^{n} a_{i,j}(-1)^{i+j}|\mathbf{A}_{i,j}|$$

and for any column j, this would be

$$|\mathbf{A}| = \sum_{i=1}^{n} a_{i,j}(-1)^{i+j}|\mathbf{A}_{i,j}|$$

Determinants

Determinants

Value of Determinants Example

$$|\mathbf{A}| = \begin{vmatrix} 1 & 4 & 3 \\ 2 & 0 & 2 \\ 1 & 3 & 5 \end{vmatrix}$$

Determinants

Value of Determinants Example

$$|\mathbf{A}| = \begin{vmatrix} 1 & 4 & 3 \\ 2 & 0 & 2 \\ 1 & 3 & 5 \end{vmatrix}$$

Expanding $|\mathbf{A}|$ by column 3, we get

$$\begin{aligned} |\mathbf{A}| &= 3(-1)^{1+3}\begin{vmatrix} 2 & 0 \\ 1 & 3 \end{vmatrix} + 2(-1)^{2+3}\begin{vmatrix} 1 & 4 \\ 1 & 3 \end{vmatrix} + 5(-1)^{3+3}\begin{vmatrix} 1 & 4 \\ 2 & 0 \end{vmatrix} \\ &= 3(6) - 2(-1) + 5(-8) = -20 \end{aligned}$$

Matrices

Value of Determinants (Contd.)

The expansion of determinants can become complex for larger orders.

Matrices

Value of Determinants (Contd.)

The expansion of determinants can become complex for larger orders. We can simplify the process by utilizing five properties.

Matrices

Value of Determinants (Contd.)

The expansion of determinants can become complex for larger orders. We can simplify the process by utilizing five properties. Note that we can interchange the words ”row” and ”column”.

Matrices

Value of Determinants (Contd.)

The expansion of determinants can become complex for larger orders. We can simplify the process by utilizing five properties. Note that we can interchange the words ”row” and ”column”.

1. If one complete row of a determinant is all zero, the value of the determinant is zero.

Matrices

Value of Determinants (Contd.)

The expansion of determinants can become complex for larger orders. We can simplify the process by utilizing five properties. Note that we can interchange the words ”row” and ”column”.

1. If one complete row of a determinant is all zero, the value of the determinant is zero.
2. If two rows have elements that are proportional to one another, the value of the determinant is zero.

Matrices

Value of Determinants (Contd.)

The expansion of determinants can become complex for larger orders. We can simplify the process by utilizing five properties. Note that we can interchange the words ”row” and ”column”.

1. If one complete row of a determinant is all zero, the value of the determinant is zero.
2. If two rows have elements that are proportional to one another, the value of the determinant is zero.
3. If two rows of a determinant are interchanged, the value of the new determinant is equal to the negative of the value of the old determinant.

Matrices

Value of Determinants (Contd.)

The expansion of determinants can become complex for larger orders. We can simplify the process by utilizing five properties. Note that we can interchange the words "row" and "column".

1. If one complete row of a determinant is all zero, the value of the determinant is zero.
2. If two rows have elements that are proportional to one another, the value of the determinant is zero.
3. If two rows of a determinant are interchanged, the value of the new determinant is equal to the negative of the value of the old determinant.
4. Elements of any row may be multiplied by a nonzero constant if the entire determinant is multiplied by the reciprocal of the constant.

Matrices

Value of Determinants (Contd.)

The expansion of determinants can become complex for larger orders. We can simplify the process by utilizing five properties. Note that we can interchange the words ”row” and ”column”.

1. If one complete row of a determinant is all zero, the value of the determinant is zero.
2. If two rows have elements that are proportional to one another, the value of the determinant is zero.
3. If two rows of a determinant are interchanged, the value of the new determinant is equal to the negative of the value of the old determinant.
4. Elements of any row may be multiplied by a nonzero constant if the entire determinant is multiplied by the reciprocal of the constant.
5. To the elements of any row, you may add a constant times the corresponding element of any other row without changing the value of the determinant.

Adjoint Matrix

Adjoint Matrix

Adjoint

If **A** is a square matrix, the **adjoint** of **A**, denoted as $\mathbf{A}^{\alpha}$, can be found using the following procedure:

Adjoint Matrix

Adjoint

If **A** is a square matrix, the **adjoint** of **A**, denoted as $\mathbf{A}^{\alpha}$, can be found using the following procedure:

1. Replace each element $a_{i,j}$ of **A** by its cofactor.

Adjoint Matrix

Adjoint

If **A** is a square matrix, the **adjoint** of **A**, denoted as $\mathbf{A}^{\alpha}$, can be found using the following procedure:

1. Replace each element $a_{i,j}$ of **A** by its cofactor.
2. Take the transpose of the matrix of cofactors found in step 1.

Adjoint Matrix

Adjoint

If **A** is a square matrix, the **adjoint** of **A**, denoted as $\mathbf{A}^{\alpha}$, can be found using the following procedure:

1. Replace each element $a_{i,j}$ of **A** by its cofactor.
2. Take the transpose of the matrix of cofactors found in step 1.
3. The resulting matrix is $\mathbf{A}^{\alpha}$, the adjoint of **A**.

Adjoint Matrix

Adjoint

If **A** is a square matrix, the **adjoint** of **A**, denoted as $\mathbf{A}^\alpha$, can be found using the following procedure:

1. Replace each element $a_{i,j}$ of **A** by its cofactor.
2. Take the transpose of the matrix of cofactors found in step 1.
3. The resulting matrix is $\mathbf{A}^\alpha$, the adjoint of **A**.

Example

Let $\gamma_{i,j} = (-1)^{i+j}|\mathbf{A}_{i,j}|$ be the cofactor for $a_{i,j}$, then

$$\mathbf{A}^\alpha = \begin{pmatrix} \gamma_{1,1} & \gamma_{2,1} & \cdots & \gamma_{n,1} \\ \gamma_{1,2} & \gamma_{2,2} & \cdots & \gamma_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{1,n} & \gamma_{2,n} & \cdots & \gamma_{n,n} \end{pmatrix}$$

Matrix Inverse

Matrix Inverse

Inverse

The **inverse** of a square matrix $\mathbf{A}$ is denoted as $\mathbf{A}^{-1}$.

Matrix Inverse

Inverse

The **inverse** of a square matrix $\mathbf{A}$ is denoted as $\mathbf{A}^{-1}$.
For a matrix to have an inverse, it must be nonsingular; i.e., its determinant cannot be zero.

Matrix Inverse

Inverse

The **inverse** of a square matrix **A** is denoted as $\mathbf{A}^{-1}$.
For a matrix to have an inverse, it must be nonsingular; i.e., its determinant cannot be zero.
Given a nonsingular matrix **A**, we find the inverse by

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\mathbf{A}^{\alpha}$$

Example

Example

Example

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 6 & 5 \end{pmatrix} \quad |\mathbf{A}| = 2(5) - 1(6) = 10 - 6 = 4$$

Example

Example

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 6 & 5 \end{pmatrix} \qquad |\mathbf{A}| = 2(5) - 1(6) = 10 - 6 = 4$$

$$\mathbf{A}^{\alpha} = \begin{pmatrix} |5| & -|1| \\ -|6| & |2| \end{pmatrix} = \begin{pmatrix} 5 & -1 \\ -6 & 2 \end{pmatrix}$$

Example

Example

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 6 & 5 \end{pmatrix} \quad |\mathbf{A}| = 2(5) - 1(6) = 10 - 6 = 4$$

$$\mathbf{A}^{\alpha} = \begin{pmatrix} |5| & -|1| \\ -|6| & |2| \end{pmatrix} = \begin{pmatrix} 5 & -1 \\ -6 & 2 \end{pmatrix}$$

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\mathbf{A}^{\alpha} = \frac{1}{4}\begin{pmatrix} 5 & -1 \\ -6 & 2 \end{pmatrix} = \begin{pmatrix} \frac{5}{4} & -\frac{1}{4} \\ -\frac{3}{2} & \frac{1}{2} \end{pmatrix}$$

Example

Example

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 6 & 5 \end{pmatrix} \quad |\mathbf{A}| = 2(5) - 1(6) = 10 - 6 = 4$$

$$\mathbf{A}^{\alpha} = \begin{pmatrix} |5| & -|1| \\ -|6| & |2| \end{pmatrix} = \begin{pmatrix} 5 & -1 \\ -6 & 2 \end{pmatrix}$$

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\mathbf{A}^{\alpha} = \frac{1}{4}\begin{pmatrix} 5 & -1 \\ -6 & 2 \end{pmatrix} = \begin{pmatrix} \frac{5}{4} & -\frac{1}{4} \\ -\frac{3}{2} & \frac{1}{2} \end{pmatrix}$$

Matrices

Matrices

Gauss-Jordan Elimination

This is another method for computing the inverse of a matrix.

Matrices

Gauss-Jordan Elimination

This is another method for computing the inverse of a matrix.
The idea is to augment the matrix with the identity matrix and then perform elementary row operations.

Matrices

Gauss-Jordan Elimination

This is another method for computing the inverse of a matrix.
The idea is to augment the matrix with the identity matrix and then perform elementary row operations.

Elementary Row Operations

1. Interchange a row i with a row j.

Matrices

Gauss-Jordan Elimination

This is another method for computing the inverse of a matrix.
The idea is to augment the matrix with the identity matrix and then perform elementary row operations.

Elementary Row Operations

1. Interchange a row *i* with a row *j*.
2. Multiply a row *i* by a nonzero scalar α.

Matrices

Gauss-Jordan Elimination

This is another method for computing the inverse of a matrix.
The idea is to augment the matrix with the identity matrix and then perform elementary row operations.

Elementary Row Operations

1. Interchange a row *i* with a row *j*.
2. Multiply a row *i* by a nonzero scalar α.
3. Replace a row *i* by a row *i* plus a multiple of some row *j*.

Matrix Rank

Matrix Rank

Rank of a Matrix

The **rank** of an m x n matrix **A**, denoted as $r(\mathbf{A})$, is the number of linearly independent columns (or rows) of **A**.

Matrix Rank

Rank of a Matrix

The **rank** of an m x n matrix **A**, denoted as $r(\mathbf{A})$, is the number of linearly independent columns (or rows) of **A**.
By definition, $r(\mathbf{A}) \leq \min\{m, n\}$.

Matrix Rank

Rank of a Matrix

The **rank** of an *m* x *n* matrix **A**, denoted as $r(\mathbf{A})$, is the number of linearly independent columns (or rows) of **A**.
By definition, $r(\mathbf{A}) \leq \min\{m, n\}$.
If $r(\mathbf{A}) = \min\{m, n\}$, then **A** is said to be of **full rank**.

Matrix Rank

Rank of a Matrix

The **rank** of an m x n matrix **A**, denoted as $r(\mathbf{A})$, is the number of linearly independent columns (or rows) of **A**.
By definition, $r(\mathbf{A}) \leq \min\{m, n\}$.
If $r(\mathbf{A}) = \min\{m, n\}$, then **A** is said to be of **full rank**.
There are several ways to get the rank, but the method used here will use elementary row operations to get

$$\left(\begin{array}{c|c} \mathbf{I}_k & \mathbf{D} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right)$$

Matrix Rank

Rank of a Matrix

The **rank** of an m x n matrix $\mathbf{A}$, denoted as $r(\mathbf{A})$, is the number of linearly independent columns (or rows) of $\mathbf{A}$.
By definition, $r(\mathbf{A}) \leq \min\{m, n\}$.
If $r(\mathbf{A}) = \min\{m, n\}$, then $\mathbf{A}$ is said to be of **full rank**.
There are several ways to get the rank, but the method used here will use elementary row operations to get

$$\left(\begin{array}{c|c} \mathbf{I}_k & \mathbf{D} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right)$$

This shows that $r(\mathbf{A}) = k$.

Example

Example

Example

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 3 & 1 \\ 2 & 1 & 2 & 3 & 0 \\ 1 & 3 & 1 & 9 & 5 \end{pmatrix}$$

Example

Example

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 3 & 1 \\ 2 & 1 & 2 & 3 & 0 \\ 1 & 3 & 1 & 9 & 5 \end{pmatrix}$$

$$\mathbf{A} = \left(\begin{array}{cc|ccc} 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 3 & 2 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right) = \left(\begin{array}{c|c} \mathbf{I}_2 & \mathbf{D} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right)$$

Example

Example

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 3 & 1 \\ 2 & 1 & 2 & 3 & 0 \\ 1 & 3 & 1 & 9 & 5 \end{pmatrix}$$

$$\mathbf{A} = \left(\begin{array}{cc|ccc} 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 3 & 2 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right) = \left(\begin{array}{c|c} \mathbf{I}_2 & \mathbf{D} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right)$$

This means that the rank of **A** is 2.

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization
6. Financial Mathematics
   - Quantitative models
   - Problem Types

Simultaneous linear Equations

Simultaneous linear Equations

Equations

One of the best known uses for matrices and determinants is for solving simultaneous linear equations.

Simultaneous linear Equations

Equations

One of the best known uses for matrices and determinants is for solving simultaneous linear equations.
Matrices and vectors give us a nice method for expressing the problem.

Simultaneous linear Equations

Equations

One of the best known uses for matrices and determinants is for solving simultaneous linear equations.
Matrices and vectors give us a nice method for expressing the problem.

Simultaneous linear Equations

Equations

One of the best known uses for matrices and determinants is for solving simultaneous linear equations.
Matrices and vectors give us a nice method for expressing the problem.

Example

$$\begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + & a_{2,n}x_n & = & b_2 \\ & & & & & & & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{array}$$

Example

Example

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}$$

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

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

Solution Set

Solutions

The set of linear equations $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ has either no solution, a unique solution, or an infinite number of solutions.

Solution Set

Solutions

The set of linear equations $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ has either no solution, a unique solution, or an infinite number of solutions.
When determining if a solution exists, we are trying to find scalars $x_1, x_2, \ldots, x_n$ so that $\mathbf{b}$ can be written as a linear combination of the columns of $\mathbf{A}$.

Solution Set

Solutions

The set of linear equations $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ has either no solution, a unique solution, or an infinite number of solutions.

When determining if a solution exists, we are trying to find scalars $x_1, x_2, \ldots, x_n$ so that $\mathbf{b}$ can be written as a linear combination of the columns of $\mathbf{A}$.

Conditions where a solutions exists for $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$:

Solution Set

Solutions

The set of linear equations $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ has either no solution, a unique solution, or an infinite number of solutions.
When determining if a solution exists, we are trying to find scalars $x_1, x_2, \ldots, x_n$ so that $\mathbf{b}$ can be written as a linear combination of the columns of $\mathbf{A}$.
Conditions where a solutions exists for $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$:

1. If $r(\mathbf{A}|\mathbf{b}) = r(\mathbf{A}) + 1$, then no solution exists.

Solution Set

Solutions

The set of linear equations $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ has either no solution, a unique solution, or an infinite number of solutions.
When determining if a solution exists, we are trying to find scalars $x_1, x_2, \ldots, x_n$ so that $\mathbf{b}$ can be written as a linear combination of the columns of $\mathbf{A}$.
Conditions where a solutions exists for $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$:

1. If $r(\mathbf{A}|\mathbf{b}) = r(\mathbf{A}) + 1$, then no solution exists.
2. If $r(\mathbf{A}|\mathbf{b}) = r(\mathbf{A})$, then there does exist a solution.

Solution Set

Solutions

The set of linear equations $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ has either no solution, a unique solution, or an infinite number of solutions.
When determining if a solution exists, we are trying to find scalars $x_1, x_2, \ldots, x_n$ so that $\mathbf{b}$ can be written as a linear combination of the columns of $\mathbf{A}$.
Conditions where a solutions exists for $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$:

1. If $r(\mathbf{A}|\mathbf{b}) = r(\mathbf{A}) + 1$, then no solution exists.
2. If $r(\mathbf{A}|\mathbf{b}) = r(\mathbf{A})$, then there does exist a solution. This is because we can write $\mathbf{b}$ as a linear combination of the columns of $\mathbf{A}$.

Solution Set

Solutions

The set of linear equations $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ has either no solution, a unique solution, or an infinite number of solutions.
When determining if a solution exists, we are trying to find scalars $x_1, x_2, \ldots, x_n$ so that $\mathbf{b}$ can be written as a linear combination of the columns of $\mathbf{A}$.
Conditions where a solutions exists for $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$:

1. If $r(\mathbf{A}|\mathbf{b}) = r(\mathbf{A}) + 1$, then no solution exists.
2. If $r(\mathbf{A}|\mathbf{b}) = r(\mathbf{A})$, then there does exist a solution. This is because we can write $\mathbf{b}$ as a linear combination of the columns of $\mathbf{A}$. Furthermore, if $r(\mathbf{A}) = n$, where n is the number of variables, then there exists a unique solution for the system of equations.

Unique solution

Unique solution

A Unique Solution of $\mathbf{Ax} = \mathbf{b}$

There are several methods for solving for a unique solution, including Cramer's rule and Gaussian elimination.

Unique solution

A Unique Solution of $\mathbf{Ax} = \mathbf{b}$

There are several methods for solving for a unique solution, including Cramer's rule and Gaussian elimination.
We will first use Cramer's rule; however, we should note that this is not an efficient approach computationally.

Unique solution

A Unique Solution of $\mathbf{Ax} = \mathbf{b}$

There are several methods for solving for a unique solution, including Cramer's rule and Gaussian elimination.
We will first use Cramer's rule; however, we should note that this is not an efficient approach computationally. Let $\mathbf{A}_j$ be the matrix $\mathbf{A}$ where the jth column is replaced by $\mathbf{b}$.

Unique solution

A Unique Solution of $\mathbf{Ax} = \mathbf{b}$

There are several methods for solving for a unique solution, including Cramer's rule and Gaussian elimination.

We will first use Cramer's rule; however, we should note that this is not an efficient approach computationally. Let $\mathbf{A}_j$ be the matrix $\mathbf{A}$ where the jth column is replaced by $\mathbf{b}$.

Cramer's rule states that the unique solution is given by $x_j = \frac{|\mathbf{A}_j|}{|\mathbf{A}|}$, for all $j = 1, \ldots, n$.

Cramer's rule

Cramer's rule

Using Cramer's Rule

$$\begin{array}{rcrcrcr} 2x_1 & + & x_2 & + & 2x_3 & = & 6 \\ 2x_1 & + & 3x_2 & + & x_3 & = & 9 \\ x_1 & + & x_2 & + & x_3 & = & 3 \end{array}$$

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 2 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 6 \\ 9 \\ 3 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Cramer's rule

Using Cramer's Rule

$$\begin{array}{rcrcrcl} 2x_1 & + & x_2 & + & 2x_3 & = & 6 \\ 2x_1 & + & 3x_2 & + & x_3 & = & 9 \\ x_1 & + & x_2 & + & x_3 & = & 3 \end{array}$$

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 2 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 6 \\ 9 \\ 3 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$x_1 = \frac{\begin{vmatrix} 6 & 1 & 2 \\ 9 & 3 & 1 \\ 3 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 1 & 2 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{vmatrix}} = \frac{6}{1} = 6 \quad x_2 = \frac{\begin{vmatrix} 2 & 6 & 2 \\ 2 & 9 & 1 \\ 1 & 3 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 1 & 2 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{vmatrix}} = 0 \quad x_3 = \frac{\begin{vmatrix} 2 & 1 & 6 \\ 2 & 3 & 9 \\ 1 & 1 & 3 \end{vmatrix}}{\begin{vmatrix} 2 & 1 & 2 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{vmatrix}} = -3$$

The Inverse Method

The Inverse Method

Using Inverses

Another approach to finding a unique solution is by using the inverse.

The Inverse Method

Using Inverses

Another approach to finding a unique solution is by using the inverse.
Given $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ and $\mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{I}$, we can see that $\mathbf{A}^{-1} \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$,

The Inverse Method

Using Inverses

Another approach to finding a unique solution is by using the inverse.
Given $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ and $\mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{I}$, we can see that $\mathbf{A}^{-1} \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$, which means that $\mathbf{I} \cdot \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$,

The Inverse Method

Using Inverses

Another approach to finding a unique solution is by using the inverse.
Given $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ and $\mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{I}$, we can see that $\mathbf{A}^{-1} \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$, which means that $\mathbf{I} \cdot \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$, and hence $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$.

The Inverse Method

Using Inverses

Another approach to finding a unique solution is by using the inverse.
Given $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ and $\mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{I}$, we can see that $\mathbf{A}^{-1} \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$, which means that $\mathbf{I} \cdot \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$, and hence $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$.

Example

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 2 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 6 \\ 9 \\ 3 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

The Inverse Method

Using Inverses

Another approach to finding a unique solution is by using the inverse.
Given $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ and $\mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{I}$, we can see that $\mathbf{A}^{-1} \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$, which means that $\mathbf{I} \cdot \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$, and hence $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$.

Example

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 2 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 6 \\ 9 \\ 3 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

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

The Inverse Method

Using Inverses

Another approach to finding a unique solution is by using the inverse.
Given $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ and $\mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{I}$, we can see that $\mathbf{A}^{-1} \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$, which means that $\mathbf{I} \cdot \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$, and hence $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$.

Example

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 2 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 6 \\ 9 \\ 3 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

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

$$\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b} = \begin{pmatrix} 2 & 1 & -5 \\ -1 & 0 & 2 \\ -1 & -1 & 4 \end{pmatrix} \begin{pmatrix} 6 \\ 9 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ -3 \end{pmatrix}$$

Linear Equations

Linear Equations

Infinite Number of Solutions

This case is one of most interest since this scenario is the most likely to happen in linear programming.

Linear Equations

Infinite Number of Solutions

This case is one of most interest since this scenario is the most likely to happen in linear programming.
This happens when $r(\mathbf{A}) = r(\mathbf{A}|\mathbf{b}) < n$, where n is the number of variables.

Linear Equations

Infinite Number of Solutions

This case is one of most interest since this scenario is the most likely to happen in linear programming.
This happens when $r(\mathbf{A}) = r(\mathbf{A}|\mathbf{b}) < n$, where n is the number of variables.

Example

$$\begin{array}{rcrcrcr} 3x_1 & + & x_2 & - & x_3 & = & 8 \\ x_1 & + & x_2 & + & x_3 & = & 4 \end{array}$$

Linear Equations

Infinite Number of Solutions

This case is one of most interest since this scenario is the most likely to happen in linear programming.
This happens when $r(\mathbf{A}) = r(\mathbf{A}|\mathbf{b}) < n$, where n is the number of variables.

Example

$$\begin{array}{rcrcrcl} 3x_1 & + & x_2 & - & x_3 & = & 8 \\ x_1 & + & x_2 & + & x_3 & = & 4 \end{array}$$

We see that $r(\mathbf{A}) = r(\mathbf{A}|\mathbf{b}) = 2 < 3$, where

$$\mathbf{A} = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$$

Linear Equations

Infinite Number of Solutions (Contd.)

For this case, we can choose r equations, where r is the rank, and find r of the variables in terms of the remaining $n - r$ variables.

Linear Equations

Infinite Number of Solutions (Contd.)

For this case, we can choose r equations, where r is the rank, and find r of the variables in terms of the remaining $n - r$ variables.

$$\begin{array}{rcrcrcr} 3x_1 & + & x_2 & - & x_3 & = & 8 \\ x_1 & + & x_2 & + & x_3 & = & 4 \end{array}$$

Linear Equations

Infinite Number of Solutions (Contd.)

For this case, we can choose r equations, where r is the rank, and find r of the variables in terms of the remaining $n - r$ variables.

$$\begin{array}{rcrcrcr} 3x_1 & + & x_2 & - & x_3 & = & 8 \\ x_1 & + & x_2 & + & x_3 & = & 4 \end{array}$$

Solving for x_1 and x_2 gets

$$\begin{array}{rcrcr} x_1 & = & 2 & + & x_3 \\ x_2 & = & 2 & - & 2x_3 \end{array}$$

Linear Equations

Infinite Number of Solutions (Contd.)

For this case, we can choose r equations, where r is the rank, and find r of the variables in terms of the remaining $n - r$ variables.

$$\begin{array}{rcrcrcr} 3x_1 & + & x_2 & - & x_3 & = & 8 \\ x_1 & + & x_2 & + & x_3 & = & 4 \end{array}$$

Solving for x_1 and x_2 gets

$$\begin{array}{rcrcr} x_1 & = & 2 & + & x_3 \\ x_2 & = & 2 & - & 2x_3 \end{array}$$

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 + x_3 \\ 2 - 2x_3 \\ x_3 \end{pmatrix}$$

Linear Equations

Infinite Number of Solutions (Contd.)

For this case, we can choose r equations, where r is the rank, and find r of the variables in terms of the remaining $n - r$ variables.

$$\begin{array}{rcrcrcl} 3x_1 & + & x_2 & - & x_3 & = & 8 \\ x_1 & + & x_2 & + & x_3 & = & 4 \end{array}$$

Solving for x_1 and x_2 gets

$$\begin{array}{rcrcr} x_1 & = & 2 & + & x_3 \\ x_2 & = & 2 & - & 2x_3 \end{array}$$

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 + x_3 \\ 2 - 2x_3 \\ x_3 \end{pmatrix}$$

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization
6. Financial Mathematics
   - Quantitative models
   - Problem Types

Sets

Sets

Definition (Convex Combination)

Sets

Definition (Convex Combination)

Given two points $\mathbf{x}$ and $\mathbf{y}$ in E^m, and $\alpha \in [0, 1]$, the parametric point $\alpha \cdot \mathbf{x} + (1 - \alpha) \cdot \mathbf{y}$ is said to be a convex combination of $\mathbf{x}$ and $\mathbf{y}$.

Sets

Definition (Convex Combination)

Given two points $\mathbf{x}$ and $\mathbf{y}$ in E^m, and $\alpha \in [0, 1]$, the parametric point $\alpha \cdot \mathbf{x} + (1 - \alpha) \cdot \mathbf{y}$ is said to be a convex combination of $\mathbf{x}$ and $\mathbf{y}$.

Note

The set of all convex combinations of $\mathbf{x}$ *and* $\mathbf{y}$ *is the line segment joining them.*

Sets

Definition (Convex Combination)

Given two points $\mathbf{x}$ and $\mathbf{y}$ in E^m, and $\alpha \in [0, 1]$, the parametric point $\alpha \cdot \mathbf{x} + (1 - \alpha) \cdot \mathbf{y}$ is said to be a convex combination of $\mathbf{x}$ and $\mathbf{y}$.

Note

The set of all convex combinations of $\mathbf{x}$ *and* $\mathbf{y}$ *is the line segment joining them.*

Definition (Convex Set)

A set S is said to be convex, if:

Sets

Definition (Convex Combination)

Given two points $\mathbf{x}$ and $\mathbf{y}$ in E^m, and $\alpha \in [0, 1]$, the parametric point $\alpha \cdot \mathbf{x} + (1 - \alpha) \cdot \mathbf{y}$ is said to be a convex combination of $\mathbf{x}$ and $\mathbf{y}$.

Note

The set of all convex combinations of $\mathbf{x}$ *and* $\mathbf{y}$ *is the line segment joining them.*

Definition (Convex Set)

A set S is said to be convex, if:
$(\forall \mathbf{x})(\forall \mathbf{y})(\forall \alpha \in [0, 1])$

Sets

Definition (Convex Combination)

Given two points $\mathbf{x}$ and $\mathbf{y}$ in E^m, and $\alpha \in [0, 1]$, the parametric point $\alpha \cdot \mathbf{x} + (1 - \alpha) \cdot \mathbf{y}$ is said to be a convex combination of $\mathbf{x}$ and $\mathbf{y}$.

Note

The set of all convex combinations of $\mathbf{x}$ *and* $\mathbf{y}$ *is the line segment joining them.*

Definition (Convex Set)

A set S is said to be convex, if:
$(\forall \mathbf{x})(\forall \mathbf{y})(\forall \alpha \in [0, 1])\ \mathbf{x}, \mathbf{y} \in S \rightarrow \alpha \cdot \mathbf{x} + (1 - \alpha) \cdot \mathbf{y} \in S.$

Sets

Definition (Convex Combination)

Given two points $\mathbf{x}$ and $\mathbf{y}$ in E^m, and $\alpha \in [0, 1]$, the parametric point $\alpha \cdot \mathbf{x} + (1 - \alpha) \cdot \mathbf{y}$ is said to be a convex combination of $\mathbf{x}$ and $\mathbf{y}$.

Note

The set of all convex combinations of $\mathbf{x}$ *and* $\mathbf{y}$ *is the line segment joining them.*

Definition (Convex Set)

A set S is said to be convex, if:
$(\forall \mathbf{x})(\forall \mathbf{y})(\forall \alpha \in [0, 1])\ \mathbf{x}, \mathbf{y} \in S \rightarrow \alpha \cdot \mathbf{x} + (1 - \alpha) \cdot \mathbf{y} \in S$.

Exercise

A set of the form $\mathbf{A} \cdot \mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}$ *is said to be a polyhedral set.*

Sets

Definition (Convex Combination)

Given two points $\mathbf{x}$ and $\mathbf{y}$ in E^m, and $\alpha \in [0, 1]$, the parametric point $\alpha \cdot \mathbf{x} + (1 - \alpha) \cdot \mathbf{y}$ is said to be a convex combination of $\mathbf{x}$ and $\mathbf{y}$.

Note

The set of all convex combinations of $\mathbf{x}$ *and* $\mathbf{y}$ *is the line segment joining them.*

Definition (Convex Set)

A set S is said to be convex, if:
$(\forall \mathbf{x})(\forall \mathbf{y})(\forall \alpha \in [0, 1])\ \mathbf{x}, \mathbf{y} \in S \rightarrow \alpha \cdot \mathbf{x} + (1 - \alpha) \cdot \mathbf{y} \in S$.

Exercise

A set of the form $\mathbf{A} \cdot \mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}$ *is said to be a polyhedral set. Argue that polyhedral sets are convex.*

Functions

Functions

Definition (Convex function)

Functions

Definition (Convex function)

Given a convex set S, a function $f : S \rightarrow \Re$ is called convex,

Functions

Definition (Convex function)

Given a convex set S, a function $f : S \rightarrow \Re$ is called convex, if $\forall\ \mathbf{x}, \mathbf{y} \in S,\ \lambda \in [0, 1]$, we have,

Functions

Definition (Convex function)

Given a convex set S, a function $f : S \rightarrow \Re$ is called convex, if $\forall\ \mathbf{x}, \mathbf{y} \in S,\ \lambda \in [0, 1]$, we have,

$$f(\lambda \cdot \mathbf{x} + (1 - \lambda) \cdot \mathbf{y}) \leq \lambda \cdot f(\mathbf{x}) + (1 - \lambda) \cdot f(\mathbf{y}).$$

Functions

Definition (Convex function)

Given a convex set S, a function $f : S \to \Re$ is called convex, if $\forall\ \mathbf{x}, \mathbf{y} \in S,\ \lambda \in [0, 1]$, we have,

$$f(\lambda \cdot \mathbf{x} + (1 - \lambda) \cdot \mathbf{y}) \leq \lambda \cdot f(\mathbf{x}) + (1 - \lambda) \cdot f(\mathbf{y}).$$

If $<$ holds as opposed to $\leq$, the function is said to be strictly convex.

Functions

Definition (Convex function)

Given a convex set S, a function $f : S \to \Re$ is called convex, if $\forall\ \mathbf{x}, \mathbf{y} \in S,\ \lambda \in [0, 1]$, we have,

$$f(\lambda \cdot \mathbf{x} + (1 - \lambda) \cdot \mathbf{y}) \leq \lambda \cdot f(\mathbf{x}) + (1 - \lambda) \cdot f(\mathbf{y}).$$

If $<$ holds as opposed to $\leq$, the function is said to be strictly convex.

Definition (Concave function)

Functions

Definition (Convex function)

Given a convex set S, a function $f : S \rightarrow \Re$ is called convex, if $\forall\ \mathbf{x}, \mathbf{y} \in S,\ \ \lambda \in [0, 1]$, we have,

$$f(\lambda \cdot \mathbf{x} + (1 - \lambda) \cdot \mathbf{y}) \leq \lambda \cdot f(\mathbf{x}) + (1 - \lambda) \cdot f(\mathbf{y}).$$

If $<$ holds as opposed to $\leq$, the function is said to be strictly convex.

Definition (Concave function)

A function f is concave if and only if $-f$ is convex.

Functions

Definition (Convex function)

Given a convex set S, a function $f : S \rightarrow \Re$ is called convex, if $\forall$ $\mathbf{x}, \mathbf{y} \in S, \ \lambda \in [0, 1]$, we have,

$$f(\lambda \cdot \mathbf{x} + (1 - \lambda) \cdot \mathbf{y}) \leq \lambda \cdot f(\mathbf{x}) + (1 - \lambda) \cdot f(\mathbf{y}).$$

If $<$ holds as opposed to $\leq$, the function is said to be strictly convex.

Definition (Concave function)

A function f is concave if and only if $-f$ is convex.

Definition

Functions

Definition (Convex function)

Given a convex set S, a function $f : S \rightarrow \Re$ is called convex, if $\forall\ \mathbf{x}, \mathbf{y} \in S,\ \ \lambda \in [0, 1]$, we have,

$$f(\lambda \cdot \mathbf{x} + (1 - \lambda) \cdot \mathbf{y}) \leq \lambda \cdot f(\mathbf{x}) + (1 - \lambda) \cdot f(\mathbf{y}).$$

If $<$ holds as opposed to $\leq$, the function is said to be strictly convex.

Definition (Concave function)

A function f is concave if and only if $-f$ is convex.

Definition

The epigraph of a function $f : S \rightarrow \Re$, is defined as the set

Functions

Definition (Convex function)

Given a convex set S, a function $f : S \rightarrow \Re$ is called convex, if $\forall\ \mathbf{x}, \mathbf{y} \in S,\ \ \lambda \in [0, 1]$, we have,

$$f(\lambda \cdot \mathbf{x} + (1 - \lambda) \cdot \mathbf{y}) \leq \lambda \cdot f(\mathbf{x}) + (1 - \lambda) \cdot f(\mathbf{y}).$$

If $<$ holds as opposed to $\leq$, the function is said to be strictly convex.

Definition (Concave function)

A function f is concave if and only if $-f$ is convex.

Definition

The epigraph of a function $f : S \rightarrow \Re$, is defined as the set $\{(\mathbf{x}, r) : \mathbf{x} \in S, f(\mathbf{x}) \leq r\}$.

Functions

Definition (Convex function)

Given a convex set S, a function $f : S \to \Re$ is called convex, if $\forall\ \mathbf{x}, \mathbf{y} \in S,\ \ \lambda \in [0, 1]$, we have,

$$f(\lambda \cdot \mathbf{x} + (1 - \lambda) \cdot \mathbf{y}) \leq \lambda \cdot f(\mathbf{x}) + (1 - \lambda) \cdot f(\mathbf{y}).$$

If $<$ holds as opposed to $\leq$, the function is said to be strictly convex.

Definition (Concave function)

A function f is concave if and only if $-f$ is convex.

Definition

The epigraph of a function $f : S \to \Re$, is defined as the set $\{(\mathbf{x}, r) : \mathbf{x} \in S, f(\mathbf{x}) \leq r\}$.

Theorem

Functions

Definition (Convex function)

Given a convex set S, a function $f : S \to \Re$ is called convex, if $\forall\ \mathbf{x}, \mathbf{y} \in S,\ \lambda \in [0, 1]$, we have,

$$f(\lambda \cdot \mathbf{x} + (1 - \lambda) \cdot \mathbf{y}) \leq \lambda \cdot f(\mathbf{x}) + (1 - \lambda) \cdot f(\mathbf{y}).$$

If $<$ holds as opposed to $\leq$, the function is said to be strictly convex.

Definition (Concave function)

A function f is concave if and only if $-f$ is convex.

Definition

The epigraph of a function $f : S \to \Re$, is defined as the set $\{(\mathbf{x}, r) : \mathbf{x} \in S, f(\mathbf{x}) \leq r\}$.

Theorem

f is a convex function if and only if its epigraph is a convex set.

Checking convexity

Checking convexity

Theorem

If f is a twice-diferentiable, univariate function, then f is convex on set S, if and only if $f''(\mathbf{x}) \geq 0$, for all $\mathbf{x} \in \mathbf{S}$.

Checking convexity

Theorem

If f is a twice-diferentiable, univariate function, then f is convex on set S, if and only if $f''(\mathbf{x}) \geq 0$, for all $\mathbf{x} \in \mathbf{S}$. A multivariate function f is convex if and only if,

Checking convexity

Theorem

If f is a twice-diferentiable, univariate function, then f is convex on set S, if and only if $f''(\mathbf{x}) \geq 0$, for all $\mathbf{x} \in \mathbf{S}$. A multivariate function f is convex if and only if, $\bigtriangledown^2 f(\mathbf{x})$ is positive semidefinite.

Checking convexity

Theorem

If f is a twice-diferentiable, univariate function, then f is convex on set S, if and only if $f''(\mathbf{x}) \geq 0$, for all $\mathbf{x} \in \mathbf{S}$. A multivariate function f is convex if and only if, $\bigtriangledown^2 f(\mathbf{x})$ is positive semidefinite. Recall that,

$$[\bigtriangledown^2 f(\mathbf{x})]_{i,j} = \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}, \ \forall i, j$$

Convex optimization theorem

Convex optimization theorem

Theorem

Convex optimization theorem

Theorem

Consider the following optimization problem:

$$\begin{aligned} & \min_{\mathbf{x}} f(x) \\ s.t. \quad & \mathbf{x} \in \mathbf{S} \end{aligned}$$

Convex optimization theorem

Theorem

Consider the following optimization problem:

$$\begin{aligned} & \min_{\mathbf{x}} f(x) \\ s.t. \quad & \mathbf{x} \in \mathbf{S} \end{aligned}$$

If S is a convex set and f is a convex function of **x** *on S, the all local optima are also global optima.*

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization
6. Financial Mathematics
   - Quantitative models
   - Problem Types

Cones

Cones

Definition

Cones

Definition

A cone is a set that is closed under positive scalar multiplication.

Cones

Definition

A cone is a set that is closed under positive scalar multiplication. It is called *pointed*, if it does not include any lines.

Cones

Definition

A cone is a set that is closed under positive scalar multiplication. It is called *pointed*, if it does not include any lines.

Note

Are cones convex?

Cones

Definition

A cone is a set that is closed under positive scalar multiplication. It is called *pointed*, if it does not include any lines.

Note

Are cones convex? We will be dealing with pointed, convex cones only.

Cone Examples

Cone Examples

Examples

Cone Examples

Examples

1. The positive orthant - $\{\mathbf{x} \in \Re^n : \mathbf{x} \geq \mathbf{0}\}$.

Cone Examples

Examples

1. The positive orthant - $\{\mathbf{x} \in \Re^n : \mathbf{x} \geq \mathbf{0}\}$.
2. Polyhedral cones - $\{\mathbf{x} \in \Re^n : \mathbf{A} \cdot \mathbf{x} \geq \mathbf{0}\}$.

Cone Examples

Examples

1. The positive orthant - $\{\mathbf{x} \in \Re^n : \mathbf{x} \geq \mathbf{0}\}$.
2. Polyhedral cones - $\{\mathbf{x} \in \Re^n : \mathbf{A} \cdot \mathbf{x} \geq \mathbf{0}\}$.
3. Lorentz cones - $\{\mathbf{x} = [x_1, \ldots x_n] \in \Re^{n+1} : x_n \geq ||(x_1, x_2 \ldots x_{n-1})||_2$.

Cone Examples

Examples

1. The positive orthant - $\{\mathbf{x} \in \Re^n : \mathbf{x} \geq \mathbf{0}\}$.
2. Polyhedral cones - $\{\mathbf{x} \in \Re^n : \mathbf{A} \cdot \mathbf{x} \geq \mathbf{0}\}$.
3. Lorentz cones - $\{\mathbf{x} = [x_1, \ldots x_n] \in \Re^{n+1} : x_n \geq ||(x_1, x_2 \ldots x_{n-1})||_2$.
4. The cone of symmetric positive semidefinite matrices - $\{\mathbf{X} \in \Re^{n \times n} : \mathbf{X} = \mathbf{X}^{\mathbf{T}}, \text{ and } \mathbf{X} \text{ is positive semidefinite}\}$.

Cone Properties

Cone Properties

Definition (Dual Cone)

If C is a cone in vector space X, with an inner product "$\cdot$", then its *dual cone* is denoted by:

Cone Properties

Definition (Dual Cone)

If C is a cone in vector space X, with an inner product "·", then its *dual cone* is denoted by:

$$C^* = \{\mathbf{x} \in X : \mathbf{x} \cdot \mathbf{y} \geq 0, \forall \mathbf{y} \in C\}.$$

Cone Properties

Definition (Dual Cone)

If C is a cone in vector space X, with an inner product "$\cdot$", then its *dual cone* is denoted by:

$$C^* = \{\mathbf{x} \in X : \mathbf{x} \cdot \mathbf{y} \geq 0, \forall \mathbf{y} \in C\}.$$

Definition (Polar Cone)

The polar cone of a cone C is the negative of its dual, i.e.,

Cone Properties

Definition (Dual Cone)

If C is a cone in vector space X, with an inner product "$\cdot$", then its *dual cone* is denoted by:

$$C^* = \{\mathbf{x} \in X : \mathbf{x} \cdot \mathbf{y} \geq 0, \forall \mathbf{y} \in C\}.$$

Definition (Polar Cone)

The polar cone of a cone C is the negative of its dual, i.e.,

$$C^P = \{\mathbf{x} \in X : \mathbf{x} \cdot \mathbf{y} \leq 0, \forall \mathbf{y} \in C\}.$$

Cone Properties

Definition (Dual Cone)

If C is a cone in vector space X, with an inner product "$\cdot$", then its *dual cone* is denoted by:

$$C^* = \{\mathbf{x} \in X : \mathbf{x} \cdot \mathbf{y} \geq 0, \forall \mathbf{y} \in C\}.$$

Definition (Polar Cone)

The polar cone of a cone C is the negative of its dual, i.e.,

$$C^P = \{\mathbf{x} \in X : \mathbf{x} \cdot \mathbf{y} \leq 0, \forall \mathbf{y} \in C\}.$$

Exercise

Show that the cone $\Re^n_+$ is its own dual cone.

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization
6. Financial Mathematics
   - Quantitative models
   - Problem Types

Sample Space and Events

Sample Space and Events

Definition

Sample Space and Events

Definition

A random experiment is an experiment whose outcome is not known in advance,

Sample Space and Events

Definition

A random experiment is an experiment whose outcome is not known in advance, but belongs to a non-empty, non-singleton set called the sample space (usually denoted by S).

Sample Space and Events

Definition

A random experiment is an experiment whose outcome is not known in advance, but belongs to a non-empty, non-singleton set called the sample space (usually denoted by S).

Example

Sample Space and Events

Definition

A random experiment is an experiment whose outcome is not known in advance, but belongs to a non-empty, non-singleton set called the sample space (usually denoted by S).

Example

(i) Suppose that the experiment consists of tossing a coin.

Sample Space and Events

Definition

A random experiment is an experiment whose outcome is not known in advance, but belongs to a non-empty, non-singleton set called the sample space (usually denoted by S).

Example

(i) Suppose that the experiment consists of tossing a coin. Then, $S = \{H, T\}$.

Sample Space and Events

Definition

A random experiment is an experiment whose outcome is not known in advance, but belongs to a non-empty, non-singleton set called the sample space (usually denoted by S).

Example

(i) Suppose that the experiment consists of tossing a coin. Then, $S = \{H, \ T\}$.
(ii) Suppose that the experiment consists of tossing a die.

Sample Space and Events

Definition

A random experiment is an experiment whose outcome is not known in advance, but belongs to a non-empty, non-singleton set called the sample space (usually denoted by S).

Example

(i) Suppose that the experiment consists of tossing a coin. Then, $S = \{H, T\}$.

(ii) Suppose that the experiment consists of tossing a die. Then, $S = \{1, 2, 3, 4, 5, 6\}$.

Sample Space and Events

Definition

A random experiment is an experiment whose outcome is not known in advance, but belongs to a non-empty, non-singleton set called the sample space (usually denoted by S).

Example

(i) Suppose that the experiment consists of tossing a coin. Then, $S = \{H, \ T\}$.

(ii) Suppose that the experiment consists of tossing a die. Then, $S = \{1, \ 2, \ 3, \ 4, \ 5, \ 6\}$.

(iii) Suppose that the experiment consists of tossing two coins.

Sample Space and Events

Definition

A random experiment is an experiment whose outcome is not known in advance, but belongs to a non-empty, non-singleton set called the sample space (usually denoted by S).

Example

(i) Suppose that the experiment consists of tossing a coin. Then, $S = \{H, T\}$.

(ii) Suppose that the experiment consists of tossing a die. Then, $S = \{1, 2, 3, 4, 5, 6\}$.

(iii) Suppose that the experiment consists of tossing two coins. Then, $S = \{HH, HT, TH, TT\}$.

Sample Space and Events

Definition

A random experiment is an experiment whose outcome is not known in advance, but belongs to a non-empty, non-singleton set called the sample space (usually denoted by S).

Example

(i) Suppose that the experiment consists of tossing a coin. Then, $S = \{H, T\}$.

(ii) Suppose that the experiment consists of tossing a die. Then, $S = \{1, 2, 3, 4, 5, 6\}$.

(iii) Suppose that the experiment consists of tossing two coins. Then, $S = \{HH, HT, TH, TT\}$.

(iv) Suppose that the experiment consists of measuring the life of a battery.

Sample Space and Events

Definition

A random experiment is an experiment whose outcome is not known in advance, but belongs to a non-empty, non-singleton set called the sample space (usually denoted by S).

Example

(i) Suppose that the experiment consists of tossing a coin. Then, $S = \{H, T\}$.

(ii) Suppose that the experiment consists of tossing a die. Then, $S = \{1, 2, 3, 4, 5, 6\}$.

(iii) Suppose that the experiment consists of tossing two coins. Then, $S = \{HH, HT, TH, TT\}$.

(iv) Suppose that the experiment consists of measuring the life of a battery. Then, $S = [0, \infty)$.

Sample Space and Events

Definition

A random experiment is an experiment whose outcome is not known in advance, but belongs to a non-empty, non-singleton set called the sample space (usually denoted by S).

Example

(i) Suppose that the experiment consists of tossing a coin. Then, $S = \{H, T\}$.

(ii) Suppose that the experiment consists of tossing a die. Then, $S = \{1, 2, 3, 4, 5, 6\}$.

(iii) Suppose that the experiment consists of tossing two coins. Then, $S = \{HH, HT, TH, TT\}$.

(iv) Suppose that the experiment consists of measuring the life of a battery. Then, $S = [0, \infty)$.

Definition

Any subset of the sample space S is called an event.

Combining Events

Combining Events

Definition

Combining Events

Definition

Given two events E and F, the event $E \cup F$ (union)

Combining Events

Definition

Given two events E and F, the event $E \cup F$ (union) is defined as the event whose outcomes are in E or F;

Combining Events

Definition

Given two events E and F, the event $E \cup F$ (union) is defined as the event whose outcomes are in E or F; e.g., in the die tossing experiment, the union of the events $E = \{2, 4\}$ and $F = \{1\}$ is $\{1, 2, 4\}$.

Combining Events

Definition

Given two events E and F, the event $E \cup F$ (union) is defined as the event whose outcomes are in E or F; e.g., in the die tossing experiment, the union of the events $E = \{2, 4\}$ and $F = \{1\}$ is $\{1, 2, 4\}$.

Definition

Combining Events

Definition

Given two events E and F, the event $E \cup F$ (union) is defined as the event whose outcomes are in E or F; e.g., in the die tossing experiment, the union of the events $E = \{2,\ 4\}$ and $F = \{1\}$ is $\{1,\ 2,\ 4\}$.

Definition

Given two events E and F, the event EF

Combining Events

Definition

Given two events E and F, the event $E \cup F$ (union) is defined as the event whose outcomes are in E or F; e.g., in the die tossing experiment, the union of the events $E = \{2, 4\}$ and $F = \{1\}$ is $\{1, 2, 4\}$.

Definition

Given two events E and F, the event EF (intersection) is defined as the event whose outcomes are in E and F;

Combining Events

Definition

Given two events E and F, the event $E \cup F$ (union) is defined as the event whose outcomes are in E or F; e.g., in the die tossing experiment, the union of the events $E = \{2, 4\}$ and $F = \{1\}$ is $\{1, 2, 4\}$.

Definition

Given two events E and F, the event EF (intersection) is defined as the event whose outcomes are in E and F; e.g., in the die tossing experiment, the intersection of the events $E = \{1, 2, 3\}$ and $F = \{1\}$ is $\{1\}$.

Combining events (contd.)

Combining events (contd.)

Definition

Combining events (contd.)

Definition

Given an event E, the event E^c (complement) denotes the event whose outcomes are in S, but not in E;

Combining events (contd.)

Definition

Given an event E, the event E^c (complement) denotes the event whose outcomes are in S, but not in E; e.g., in the die tossing experiment, the complement of the event $E = \{1,\ 2,\ 3\}$ is $\{4,\ 5,\ 6\}$.

Combining events (contd.)

Definition

Given an event E, the event E^c (complement) denotes the event whose outcomes are in S, but not in E; e.g., in the die tossing experiment, the complement of the event $E = \{1, 2, 3\}$ is $\{4, 5, 6\}$.

Definition

Combining events (contd.)

Definition

Given an event E, the event E^c (complement) denotes the event whose outcomes are in S, but not in E; e.g., in the die tossing experiment, the complement of the event $E = \{1, 2, 3\}$ is $\{4, 5, 6\}$.

Definition

If events E and F have no outcomes in common, then $EF = \emptyset$ and

Combining events (contd.)

Definition

Given an event E, the event E^c (complement) denotes the event whose outcomes are in S, but not in E; e.g., in the die tossing experiment, the complement of the event $E = \{1,\ 2,\ 3\}$ is $\{4,\ 5,\ 6\}$.

Definition

If events E and F have no outcomes in common, then $EF = \emptyset$ and E and F are said to be *mutually exclusive*.

Combining events (contd.)

Definition

Given an event E, the event E^c (complement) denotes the event whose outcomes are in S, but not in E; e.g., in the die tossing experiment, the complement of the event $E = \{1, 2, 3\}$ is $\{4, 5, 6\}$.

Definition

If events E and F have no outcomes in common, then $EF = \emptyset$ and E and F are said to be *mutually exclusive*. In this case, $P(EF) = 0$;

Combining events (contd.)

Definition

Given an event E, the event E^c (complement) denotes the event whose outcomes are in S, but not in E; e.g., in the die tossing experiment, the complement of the event $E = \{1, 2, 3\}$ is $\{4, 5, 6\}$.

Definition

If events E and F have no outcomes in common, then $EF = \emptyset$ and E and F are said to be *mutually exclusive*. In this case, $P(EF) = 0$; in the single coin tossing experiment the events $\{H\}$ and $\{T\}$ are mutually exclusive.

Combining events (contd.)

Definition

Given an event E, the event E^c (complement) denotes the event whose outcomes are in S, but not in E; e.g., in the die tossing experiment, the complement of the event $E = \{1, 2, 3\}$ is $\{4, 5, 6\}$.

Definition

If events E and F have no outcomes in common, then $EF = \emptyset$ and E and F are said to be *mutually exclusive*. In this case, $P(EF) = 0$; in the single coin tossing experiment the events $\{H\}$ and $\{T\}$ are mutually exclusive.

Note

Never forget that events are sets.

Combining events (contd.)

Definition

Given an event E, the event E^c (complement) denotes the event whose outcomes are in S, but not in E; e.g., in the die tossing experiment, the complement of the event $E = \{1, 2, 3\}$ is $\{4, 5, 6\}$.

Definition

If events E and F have no outcomes in common, then $EF = \emptyset$ and E and F are said to be *mutually exclusive*. In this case, $P(EF) = 0$; in the single coin tossing experiment the events $\{H\}$ and $\{T\}$ are mutually exclusive.

Note

Never forget that events are sets. This is particularly important when using logic to reason about them.

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization
6. Financial Mathematics
   - Quantitative models
   - Problem Types

Defining Probabilities on Events

Defining Probabilities on Events

Assigning probabilities

Defining Probabilities on Events

Assigning probabilities

Let S denote a sample space.

Defining Probabilities on Events

Assigning probabilities

Let S denote a sample space. We assume that the number $P(E)$ is assigned to each event E in S, such that:

Defining Probabilities on Events

Assigning probabilities

Let S denote a sample space. We assume that the number $P(E)$ is assigned to each event E in S, such that:

(i) $0 \leq P(E) \leq 1$.

Defining Probabilities on Events

Assigning probabilities

Let S denote a sample space. We assume that the number $P(E)$ is assigned to each event E in S, such that:

(i) $0 \le P(E) \le 1$.

(ii) $P(S) = 1$.

Defining Probabilities on Events

Assigning probabilities

Let S denote a sample space. We assume that the number $P(E)$ is assigned to each event E in S, such that:

(i) $0 \leq P(E) \leq 1$.

(ii) $P(S) = 1$.

(iii) If $E_1, E_2, \ldots, E_n$ are mutually exclusive events, then,

Defining Probabilities on Events

Assigning probabilities

Let S denote a sample space. We assume that the number $P(E)$ is assigned to each event E in S, such that:

(i) $0 \leq P(E) \leq 1$.

(ii) $P(S) = 1$.

(iii) If $E_1, E_2, \ldots, E_n$ are mutually exclusive events, then,

$$P(E_1 \cup E_2 \ldots E_n) = \sum_{i=1}^{n} P(E_i).$$

Defining Probabilities on Events

Assigning probabilities

Let S denote a sample space. We assume that the number $P(E)$ is assigned to each event E in S, such that:

(i) $0 \leq P(E) \leq 1$.

(ii) $P(S) = 1$.

(iii) If $E_1, E_2, \ldots, E_n$ are mutually exclusive events, then,

$$P(E_1 \cup E_2 \ldots E_n) = \sum_{i=1}^{n} P(E_i).$$

$P(E)$ is called the probability of event E.

Defining Probabilities on Events

Assigning probabilities

Let S denote a sample space. We assume that the number $P(E)$ is assigned to each event E in S, such that:

(i) $0 \leq P(E) \leq 1$.

(ii) $P(S) = 1$.

(iii) If $E_1, E_2, \ldots, E_n$ are mutually exclusive events, then,

$$P(E_1 \cup E_2 \ldots E_n) = \sum_{i=1}^{n} P(E_i).$$

$P(E)$ is called the probability of event E. The 2-tuple (S, P) is called a probability space.

Defining Probabilities on Events

Assigning probabilities

Let S denote a sample space. We assume that the number $P(E)$ is assigned to each event E in S, such that:

(i) $0 \leq P(E) \leq 1$.

(ii) $P(S) = 1$.

(iii) If $E_1, E_2, \ldots, E_n$ are mutually exclusive events, then,

$$P(E_1 \cup E_2 \ldots E_n) = \sum_{i=1}^{n} P(E_i).$$

$P(E)$ is called the probability of event E. The 2-tuple (S, P) is called a probability space. The above three conditions are called the axioms of probability theory.

Two Identities

Two Identities

Note

Two Identities

Note

(i) *Let E be an arbitrary event on the sample space S.*

Two Identities

Note

(i) *Let E be an arbitrary event on the sample space S. Then,* $P(E) + P(E^c) = 1$.

Two Identities

Note

(i) *Let E be an arbitrary event on the sample space S. Then, $P(E) + P(E^c) = 1$.*

(ii) *Let E and F denote two arbitrary events on the sample space S.*

Two Identities

Note

(i) *Let E be an arbitrary event on the sample space S. Then,* $P(E) + P(E^c) = 1$.

(ii) *Let E and F denote two arbitrary events on the sample space S. Then,* $P(E \cup F) = P(E) + P(F) - P(EF)$.

Two Identities

Note

(i) *Let E be an arbitrary event on the sample space S. Then, $P(E) + P(E^c) = 1$.*

(ii) *Let E and F denote two arbitrary events on the sample space S. Then, $P(E \cup F) = P(E) + P(F) - P(EF)$. What is $P(E \cup F)$, when E and F are mutually exclusive?*

Two Identities

Note

(i) *Let E be an arbitrary event on the sample space S. Then, $P(E) + P(E^c) = 1$.*

(ii) *Let E and F denote two arbitrary events on the sample space S. Then,*
$P(E \cup F) = P(E) + P(F) - P(EF)$.
What is $P(E \cup F)$, when E and F are mutually exclusive?
Let G be another event on S. What is $P(E \cup F \cup G)$?

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization
6. Financial Mathematics
   - Quantitative models
   - Problem Types

Conditional Probability

Conditional Probability

Motivation

Conditional Probability

Motivation

Consider the experiment of tossing two fair coins.

Conditional Probability

Motivation

Consider the experiment of tossing two fair coins. What is the probability that both coins turn up heads?

Conditional Probability

Motivation

Consider the experiment of tossing two fair coins. What is the probability that both coins turn up heads? Now, assume that the first coin turns up heads.

Conditional Probability

Motivation

Consider the experiment of tossing two fair coins. What is the probability that both coins turn up heads? Now, assume that the first coin turns up heads. What is the probability that both coins turn up heads?

Conditional Probability

Motivation

Consider the experiment of tossing two fair coins. What is the probability that both coins turn up heads? Now, assume that the first coin turns up heads. What is the probability that both coins turn up heads?

Definition

Conditional Probability

Motivation

Consider the experiment of tossing two fair coins. What is the probability that both coins turn up heads? Now, assume that the first coin turns up heads. What is the probability that both coins turn up heads?

Definition

Let E and F denote two events on a sample space S. The conditional probability of E, given that the event F has occurred is denoted by $P(E \mid F)$

Conditional Probability

Motivation

Consider the experiment of tossing two fair coins. What is the probability that both coins turn up heads? Now, assume that the first coin turns up heads. What is the probability that both coins turn up heads?

Definition

Let E and F denote two events on a sample space S. The conditional probability of E, given that the event F has occurred is denoted by $P(E \mid F)$ and is equal to $\frac{P(EF)}{P(F)}$, assuming $P(F) > 0$.

Conditional Probability

Motivation

Consider the experiment of tossing two fair coins. What is the probability that both coins turn up heads? Now, assume that the first coin turns up heads. What is the probability that both coins turn up heads?

Definition

Let E and F denote two events on a sample space S. The conditional probability of E, given that the event F has occurred is denoted by $P(E \mid F)$ and is equal to $\frac{P(EF)}{P(F)}$, assuming $P(F) > 0$.

Example

Conditional Probability

Motivation

Consider the experiment of tossing two fair coins. What is the probability that both coins turn up heads? Now, assume that the first coin turns up heads. What is the probability that both coins turn up heads?

Definition

Let E and F denote two events on a sample space S. The conditional probability of E, given that the event F has occurred is denoted by $P(E \mid F)$ and is equal to $\frac{P(EF)}{P(F)}$, assuming $P(F) > 0$.

Example

In the previously discussed coin tossing example, let E denote the event that both coins turn up heads and F denote the event that the first coin turns up heads.

Conditional Probability

Motivation

Consider the experiment of tossing two fair coins. What is the probability that both coins turn up heads? Now, assume that the first coin turns up heads. What is the probability that both coins turn up heads?

Definition

Let E and F denote two events on a sample space S. The conditional probability of E, given that the event F has occurred is denoted by $P(E \mid F)$ and is equal to $\frac{P(EF)}{P(F)}$, assuming $P(F) > 0$.

Example

In the previously discussed coin tossing example, let E denote the event that both coins turn up heads and F denote the event that the first coin turns up heads. Accordingly, we are interested in $P(E \mid F)$.

Conditional Probability

Motivation

Consider the experiment of tossing two fair coins. What is the probability that both coins turn up heads? Now, assume that the first coin turns up heads. What is the probability that both coins turn up heads?

Definition

Let E and F denote two events on a sample space S. The conditional probability of E, given that the event F has occurred is denoted by $P(E \mid F)$ and is equal to $\frac{P(EF)}{P(F)}$, assuming $P(F) > 0$.

Example

In the previously discussed coin tossing example, let E denote the event that both coins turn up heads and F denote the event that the first coin turns up heads. Accordingly, we are interested in $P(E \mid F)$. Observe that $P(F) = \frac{1}{2}$ and $P(EF) = \frac{1}{4}$.

Conditional Probability

Motivation

Consider the experiment of tossing two fair coins. What is the probability that both coins turn up heads? Now, assume that the first coin turns up heads. What is the probability that both coins turn up heads?

Definition

Let E and F denote two events on a sample space S. The conditional probability of E, given that the event F has occurred is denoted by $P(E \mid F)$ and is equal to $\frac{P(EF)}{P(F)}$, assuming $P(F) > 0$.

Example

In the previously discussed coin tossing example, let E denote the event that both coins turn up heads and F denote the event that the first coin turns up heads. Accordingly, we are interested in $P(E \mid F)$. Observe that $P(F) = \frac{1}{2}$ and $P(EF) = \frac{1}{4}$. Hence, $P(E \mid F) = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$.

Conditional Probability

Motivation

Consider the experiment of tossing two fair coins. What is the probability that both coins turn up heads? Now, assume that the first coin turns up heads. What is the probability that both coins turn up heads?

Definition

Let E and F denote two events on a sample space S. The conditional probability of E, given that the event F has occurred is denoted by $P(E \mid F)$ and is equal to $\frac{P(EF)}{P(F)}$, assuming $P(F) > 0$.

Example

In the previously discussed coin tossing example, let E denote the event that both coins turn up heads and F denote the event that the first coin turns up heads. Accordingly, we are interested in $P(E \mid F)$. Observe that $P(F) = \frac{1}{2}$ and $P(EF) = \frac{1}{4}$. Hence, $P(E \mid F) = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$. Notice that $P(E) = \frac{1}{4} \neq P(E \mid F)$.

Independent Events

Independent Events

Definition

Independent Events

Definition

Two events E and F on a sample space S are said to be independent, if the occurrence of one does not affect the occurrence of the other.

Independent Events

Definition

Two events E and F on a sample space S are said to be independent, if the occurrence of one does not affect the occurrence of the other.
Mathematically,

Independent Events

Definition

Two events E and F on a sample space S are said to be independent, if the occurrence of one does not affect the occurrence of the other.
Mathematically,

$$P(E \mid F) = P(E).$$

Independent Events

Definition

Two events E and F on a sample space S are said to be independent, if the occurrence of one does not affect the occurrence of the other.
Mathematically,

$$P(E \mid F) = P(E).$$

Alternatively,

$$P(EF) = P(E) \cdot P(F)$$

Independent Events

Definition

Two events E and F on a sample space S are said to be independent, if the occurrence of one does not affect the occurrence of the other.
Mathematically,

$$P(E \mid F) = P(E).$$

Alternatively,

$$P(EF) = P(E) \cdot P(F)$$

Exercise

Independent Events

Definition

Two events E and F on a sample space S are said to be independent, if the occurrence of one does not affect the occurrence of the other. Mathematically,

$$P(E \mid F) = P(E).$$

Alternatively,

$$P(EF) = P(E) \cdot P(F)$$

Exercise

Consider the experiment of tossing two fair dice.

Independent Events

Definition

Two events E and F on a sample space S are said to be independent, if the occurrence of one does not affect the occurrence of the other.
Mathematically,

$$P(E \mid F) = P(E).$$

Alternatively,

$$P(EF) = P(E) \cdot P(F)$$

Exercise

Consider the experiment of tossing two fair dice. Let F denote the event that the first die turns up 4*.*

Independent Events

Definition

Two events E and F on a sample space S are said to be independent, if the occurrence of one does not affect the occurrence of the other.
Mathematically,

$$P(E \mid F) = P(E).$$

Alternatively,

$$P(EF) = P(E) \cdot P(F)$$

Exercise

Consider the experiment of tossing two fair dice. Let F denote the event that the first die turns up 4. Let E_1 denote the event that the sum of the faces of the two dice is 6.

Independent Events

Definition

Two events E and F on a sample space S are said to be independent, if the occurrence of one does not affect the occurrence of the other.
Mathematically,

$$P(E \mid F) = P(E).$$

Alternatively,

$$P(EF) = P(E) \cdot P(F)$$

Exercise

Consider the experiment of tossing two fair dice. Let F denote the event that the first die turns up 4. Let E_1 denote the event that the sum of the faces of the two dice is 6. Let E_2 denote the event that the sum of the faces of the two dice is 7.

Independent Events

Definition

Two events E and F on a sample space S are said to be independent, if the occurrence of one does not affect the occurrence of the other.
Mathematically,

$$P(E \mid F) = P(E).$$

Alternatively,

$$P(EF) = P(E) \cdot P(F)$$

Exercise

Consider the experiment of tossing two fair dice. Let F denote the event that the first die turns up 4. Let E_1 denote the event that the sum of the faces of the two dice is 6. Let E_2 denote the event that the sum of the faces of the two dice is 7. Are E_1 and F independent?

Independent Events

Definition

Two events E and F on a sample space S are said to be independent, if the occurrence of one does not affect the occurrence of the other.
Mathematically,

$$P(E \mid F) = P(E).$$

Alternatively,

$$P(EF) = P(E) \cdot P(F)$$

Exercise

Consider the experiment of tossing two fair dice. Let F denote the event that the first die turns up 4. Let E_1 denote the event that the sum of the faces of the two dice is 6. Let E_2 denote the event that the sum of the faces of the two dice is 7. Are E_1 and F independent? How about E_2 and F?

Bayes' Formula

Bayes' Formula

Derivation

Bayes' Formula

Derivation

Let E and F denote two arbitrary events on a sample space S.

Bayes' Formula

Derivation

Let E and F denote two arbitrary events on a sample space S. Clearly, $E = EF \cup EF^c$, where the events EF and EF^c are mutually exclusive.

Bayes' Formula

Derivation

Let E and F denote two arbitrary events on a sample space S. Clearly, $E = EF \cup EF^c$, where the events EF and EF^c are mutually exclusive. Now, observe that,

Bayes' Formula

Derivation

Let E and F denote two arbitrary events on a sample space S. Clearly, $E = EF \cup EF^c$, where the events EF and EF^c are mutually exclusive. Now, observe that,

$$P(E) =$$

Bayes' Formula

Derivation

Let E and F denote two arbitrary events on a sample space S. Clearly, $E = EF \cup EF^c$, where the events EF and EF^c are mutually exclusive. Now, observe that,

$$P(E) \quad = \quad P(EF) + P(EF^c)$$

Bayes' Formula

Derivation

Let E and F denote two arbitrary events on a sample space S. Clearly, $E = EF \cup EF^c$, where the events EF and EF^c are mutually exclusive. Now, observe that,

$$\begin{aligned} P(E) &= P(EF) + P(EF^c) \\ &= P(E \mid F)P(F) + P(E \mid F^c)P(F^c) \end{aligned}$$

Bayes' Formula

Derivation

Let E and F denote two arbitrary events on a sample space S. Clearly, $E = EF \cup EF^c$, where the events EF and EF^c are mutually exclusive. Now, observe that,

$$\begin{aligned} P(E) &= P(EF) + P(EF^c) \\ &= P(E \mid F)P(F) + P(E \mid F^c)P(F^c) \\ &= P(E \mid F)P(F) + P(E \mid F^c)(1 - P(F)) \end{aligned}$$

Bayes' Formula

Derivation

Let E and F denote two arbitrary events on a sample space S. Clearly, $E = EF \cup EF^c$, where the events EF and EF^c are mutually exclusive. Now, observe that,

$$\begin{aligned} P(E) &= P(EF) + P(EF^c) \\ &= P(E \mid F)P(F) + P(E \mid F^c)P(F^c) \\ &= P(E \mid F)P(F) + P(E \mid F^c)(1 - P(F)) \end{aligned}$$

Thus, the probability of an event E

Bayes' Formula

Derivation

Let E and F denote two arbitrary events on a sample space S. Clearly, $E = EF \cup EF^c$, where the events EF and EF^c are mutually exclusive. Now, observe that,

$$\begin{aligned} P(E) &= P(EF) + P(EF^c) \\ &= P(E \mid F)P(F) + P(E \mid F^c)P(F^c) \\ &= P(E \mid F)P(F) + P(E \mid F^c)(1 - P(F)) \end{aligned}$$

Thus, the probability of an event E is the weighted average of the conditional probability of E, given that event F has occurred and the conditional probability of E, given that event F has not occurred,

Bayes' Formula

Derivation

Let E and F denote two arbitrary events on a sample space S. Clearly, $E = EF \cup EF^c$, where the events EF and EF^c are mutually exclusive. Now, observe that,

$$\begin{aligned} P(E) &= P(EF) + P(EF^c) \\ &= P(E \mid F)P(F) + P(E \mid F^c)P(F^c) \\ &= P(E \mid F)P(F) + P(E \mid F^c)(1 - P(F)) \end{aligned}$$

Thus, the probability of an event E is the weighted average of the conditional probability of E, given that event F has occurred and the conditional probability of E, given that event F has not occurred, each conditional probability being given as much weight as the probability of the event that it is conditioned on, has of occurring.

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization
6. Financial Mathematics
   - Quantitative models
   - Problem Types

Random Variables

Random Variables

Motivation

Random Variables

Motivation

In case of certain random experiments, we are not so much interested in the actual outcome,

Random Variables

Motivation

In case of certain random experiments, we are not so much interested in the actual outcome, but in some function of the outcome, e.g.,

Random Variables

Motivation

In case of certain random experiments, we are not so much interested in the actual outcome, but in some function of the outcome, e.g., in the experiment of tossing two dice, we could be interested in knowing whether or not the the sum of the upturned faces is 7.

Random Variables

Motivation

In case of certain random experiments, we are not so much interested in the actual outcome, but in some function of the outcome, e.g., in the experiment of tossing two dice, we could be interested in knowing whether or not the the sum of the upturned faces is 7. We may not care whether the actual outcome is $(1, 6)$, $(6, 1)$, or $\ldots$.

Random Variables

Motivation

In case of certain random experiments, we are not so much interested in the actual outcome, but in some function of the outcome, e.g., in the experiment of tossing two dice, we could be interested in knowing whether or not the the sum of the upturned faces is 7. We may not care whether the actual outcome is $(1, 6)$, $(6, 1)$, or $\ldots$.

Example

Random Variables

Motivation

In case of certain random experiments, we are not so much interested in the actual outcome, but in some function of the outcome, e.g., in the experiment of tossing two dice, we could be interested in knowing whether or not the the sum of the upturned faces is 7. We may not care whether the actual outcome is $(1, 6)$, $(6, 1)$, or ....

Example

Let X denote the random variable that is defined as the sum of two fair dice.

Random Variables

Motivation

In case of certain random experiments, we are not so much interested in the actual outcome, but in some function of the outcome, e.g., in the experiment of tossing two dice, we could be interested in knowing whether or not the the sum of the upturned faces is 7. We may not care whether the actual outcome is $(1, 6)$, $(6, 1)$, or $\ldots$.

Example

Let X denote the random variable that is defined as the sum of two fair dice. What are the values that X can take?

Random Variables

Motivation

In case of certain random experiments, we are not so much interested in the actual outcome, but in some function of the outcome, e.g., in the experiment of tossing two dice, we could be interested in knowing whether or not the the sum of the upturned faces is 7. We may not care whether the actual outcome is $(1, 6)$, $(6, 1)$, or $\ldots$.

Example

Let X denote the random variable that is defined as the sum of two fair dice. What are the values that X can take?

$$P\{X = 1\} \quad =$$

Random Variables

Motivation

In case of certain random experiments, we are not so much interested in the actual outcome, but in some function of the outcome, e.g., in the experiment of tossing two dice, we could be interested in knowing whether or not the the sum of the upturned faces is 7. We may not care whether the actual outcome is $(1,6)$, $(6,1)$, or $\ldots$.

Example

Let X denote the random variable that is defined as the sum of two fair dice. What are the values that X can take?

$$P\{X = 1\} = 0$$

Random Variables

Motivation

In case of certain random experiments, we are not so much interested in the actual outcome, but in some function of the outcome, e.g., in the experiment of tossing two dice, we could be interested in knowing whether or not the the sum of the upturned faces is 7. We may not care whether the actual outcome is $(1, 6)$, $(6, 1)$, or $\ldots$.

Example

Let X denote the random variable that is defined as the sum of two fair dice. What are the values that X can take?

$$
\begin{aligned}
P\{X = 1\} &= 0 \\
P\{X = 2\} &= \frac{1}{36}
\end{aligned}
$$

Random Variables

Motivation

In case of certain random experiments, we are not so much interested in the actual outcome, but in some function of the outcome, e.g., in the experiment of tossing two dice, we could be interested in knowing whether or not the the sum of the upturned faces is 7. We may not care whether the actual outcome is $(1, 6)$, $(6, 1)$, or ....

Example

Let X denote the random variable that is defined as the sum of two fair dice. What are the values that X can take?

$$\begin{aligned} P\{X = 1\} &= 0 \\ P\{X = 2\} &= \frac{1}{36} \\ &\vdots \end{aligned}$$

Random Variables

Motivation

In case of certain random experiments, we are not so much interested in the actual outcome, but in some function of the outcome, e.g., in the experiment of tossing two dice, we could be interested in knowing whether or not the the sum of the upturned faces is 7. We may not care whether the actual outcome is $(1, 6)$, $(6, 1)$, or ....

Example

Let X denote the random variable that is defined as the sum of two fair dice. What are the values that X can take?

$$P\{X = 1\} = 0$$

$$P\{X = 2\} = \frac{1}{36}$$

$$\vdots$$

$$P\{X = 12\} = \frac{1}{36}$$

The Bernoulli Random Variable

The Bernoulli Random Variable

Main idea

The Bernoulli Random Variable

Main idea

Consider an experiment which has exactly two outcomes;

The Bernoulli Random Variable

Main idea

Consider an experiment which has exactly two outcomes; one is labeled a “success” and the other a “failure”.

The Bernoulli Random Variable

Main idea

Consider an experiment which has exactly two outcomes; one is labeled a “success” and the other a “failure”.

If we let the random variable X assume the value 1, if the experiment was a success and 0, if the experiment was a failure, then X is said to be a Bernoulli random variable.

The Bernoulli Random Variable

Main idea

Consider an experiment which has exactly two outcomes; one is labeled a "success" and the other a "failure".

If we let the random variable X assume the value 1, if the experiment was a success and 0, if the experiment was a failure, then X is said to be a Bernoulli random variable.

Assume that the probability that the experiment results in a success is p.

The Bernoulli Random Variable

Main idea

Consider an experiment which has exactly two outcomes; one is labeled a “success” and the other a “failure”.

If we let the random variable X assume the value 1, if the experiment was a success and 0, if the experiment was a failure, then X is said to be a Bernoulli random variable.

Assume that the probability that the experiment results in a success is p.

The probability mass function of X is given by:

The Bernoulli Random Variable

Main idea

Consider an experiment which has exactly two outcomes; one is labeled a "success" and the other a "failure".

If we let the random variable X assume the value 1, if the experiment was a success and 0, if the experiment was a failure, then X is said to be a Bernoulli random variable.

Assume that the probability that the experiment results in a success is p.

The probability mass function of X is given by:

$$p(1) \quad = \quad P\{X = 1\} = p$$

The Bernoulli Random Variable

Main idea

Consider an experiment which has exactly two outcomes; one is labeled a “success” and the other a “failure”.

If we let the random variable X assume the value 1, if the experiment was a success and 0, if the experiment was a failure, then X is said to be a Bernoulli random variable.

Assume that the probability that the experiment results in a success is p.

The probability mass function of X is given by:

$$
\begin{aligned}
p(1) &= P\{X = 1\} = p \\
p(0) &= P\{X = 0\} = 1 - p.
\end{aligned}
$$

The Binomial Random Variable

The Binomial Random Variable

Motivation

The Binomial Random Variable

Motivation

Consider an experiment which consists of n independent Bernoulli trials, with the probability of success in each trial being p.

The Binomial Random Variable

Motivation

Consider an experiment which consists of n independent Bernoulli trials, with the probability of success in each trial being p.

If X is the random variable that counts the number of successes in the n trials, then X is said to be a Binomial Random Variable.

The Binomial Random Variable

Motivation

Consider an experiment which consists of n independent Bernoulli trials, with the probability of success in each trial being p.

If X is the random variable that counts the number of successes in the n trials, then X is said to be a Binomial Random Variable.

The probability mass function of X is given by:

The Binomial Random Variable

Motivation

Consider an experiment which consists of n independent Bernoulli trials, with the probability of success in each trial being p.

If X is the random variable that counts the number of successes in the n trials, then X is said to be a Binomial Random Variable.

The probability mass function of X is given by:

$$p(i) = P\{X = i\} =$$

The Binomial Random Variable

Motivation

Consider an experiment which consists of n independent Bernoulli trials, with the probability of success in each trial being p.

If X is the random variable that counts the number of successes in the n trials, then X is said to be a Binomial Random Variable.

The probability mass function of X is given by:

$$p(i) = P\{X = i\} = C(n, i) \cdot p^i \cdot (1-p)^{n-i}, \; i = 0, 1, 2, \ldots n$$

The Geometric Random Variable

The Geometric Random Variable

Motivation

The Geometric Random Variable

Motivation

Suppose that independent Bernoulli trials, each with probability p of success are performed until a success occurs.

The Geometric Random Variable

Motivation

Suppose that independent Bernoulli trials, each with probability p of success are performed until a success occurs.

If X is the random variable that counts the number of trials until the first success, then X is said to be a geometric random variable.

The Geometric Random Variable

Motivation

Suppose that independent Bernoulli trials, each with probability p of success are performed until a success occurs.

If X is the random variable that counts the number of trials until the first success, then X is said to be a geometric random variable.

The probability mass function of X is given by:

The Geometric Random Variable

Motivation

Suppose that independent Bernoulli trials, each with probability p of success are performed until a success occurs.

If X is the random variable that counts the number of trials until the first success, then X is said to be a geometric random variable.

The probability mass function of X is given by:

$$p(i) = P\{X = i\} =$$

The Geometric Random Variable

Motivation

Suppose that independent Bernoulli trials, each with probability p of success are performed until a success occurs.

If X is the random variable that counts the number of trials until the first success, then X is said to be a geometric random variable.

The probability mass function of X is given by:

$$p(i) = P\{X = i\} = (1 - p)^{i-1} \cdot p, \; i = 1, 2, \ldots$$

Features of a random variable

Features of a random variable

Features

Features of a random variable

Features

Associated with each random variable are the following parameters:

Features of a random variable

Features

Associated with each random variable are the following parameters:

1. Probability mass function (pmt)

Features of a random variable

Features

Associated with each random variable are the following parameters:

1. Probability mass function (pmt) (Already discussed).

Features of a random variable

Features

Associated with each random variable are the following parameters:

1. Probability mass function (pmt) (Already discussed).
2. Cumulative distribution function or distribution function.

Features of a random variable

Features

Associated with each random variable are the following parameters:

1. Probability mass function (pmt) (Already discussed).
2. Cumulative distribution function or distribution function.
3. Expectation.

Features of a random variable

Features

Associated with each random variable are the following parameters:

1. Probability mass function (pmt) (Already discussed).
2. Cumulative distribution function or distribution function.
3. Expectation.
4. Variance.

Distribution Function

Distribution Function

Definition (Distribution Function)

Distribution Function

Definition (Distribution Function)

For a random variable X, the distribution function $F(\cdot)$ is defined for any real number b, $-\infty < b < \infty$, by

$$F(b) = P(X \leq b).$$

Expectation

Expectation

Definition (Expectation)

Expectation

Definition (Expectation)

Let X denote a discrete random variable with probability mass function $p(x)$.

Expectation

Definition (Expectation)

Let X denote a discrete random variable with probability mass function $p(x)$. The expected value of X, denoted by $E[X]$ is defined by:

$$E[X] = \sum_{x} x \cdot p(x).$$

Expectation

Definition (Expectation)

Let X denote a discrete random variable with probability mass function $p(x)$. The expected value of X, denoted by $E[X]$ is defined by:

$$E[X] = \sum_{x} x \cdot p(x).$$

Note

Expectation

Definition (Expectation)

Let X denote a discrete random variable with probability mass function $p(x)$. The expected value of X, denoted by $E[X]$ is defined by:

$$E[X] = \sum_{x} x \cdot p(x).$$

Note

$E[X]$ is the weighted average of the possible values that X can assume,

Expectation

Definition (Expectation)

Let X denote a discrete random variable with probability mass function $p(x)$. The expected value of X, denoted by $E[X]$ is defined by:

$$E[X] = \sum_{x} x \cdot p(x).$$

Note

$E[X]$ is the weighted average of the possible values that X can assume, each value being weighted by the probability that X assumes that value.

Variance and Covariance

Variance and Covariance

Definition (Variance)

The variance of a random variable X i(denoted by $Var(X)$ or σ^2) is given by

Variance and Covariance

Definition (Variance)

The variance of a random variable X i(denoted by $Var(X)$ or σ^2) is given by

$$E[(X - E[X])^2].$$

Variance and Covariance

Definition (Variance)

The variance of a random variable X i(denoted by $Var(X)$ or σ^2) is given by

$$E[(X - E[X])^2].$$

Definition (Covariance)

Given two (jointly distributed) random variables X and Y, the covariance between X and Y is defined as:

Variance and Covariance

Definition (Variance)

The variance of a random variable X i(denoted by $Var(X)$ or σ^2) is given by

$$E[(X - E[X])^2].$$

Definition (Covariance)

Given two (jointly distributed) random variables X and Y, the covariance between X and Y is defined as:

$$Cov(X, Y) = E[(X - E(X)) \cdot (Y - E(Y)).$$

Parameters of the important Random Variables

Parameters of the important Random Variables

Parameter table

Variable type	Expectation	Variance
Bernoulli	p	$p \cdot (1-p)$
Binomial	$n \cdot p$	$n \cdot p \cdot (1-p)$
Geometric	$\frac{1}{p}$	$\frac{1-p}{p^2}$

Parameters of the important Random Variables

Parameter table

Variable type	Expectation	Variance
Bernoulli	p	$p \cdot (1-p)$
Binomial	$n \cdot p$	$n \cdot p \cdot (1-p)$
Geometric	$\frac{1}{p}$	$\frac{1-p}{p^2}$

Exercise

Parameters of the important Random Variables

Parameter table

Variable type	Expectation	Variance
Bernoulli	p	$p \cdot (1-p)$
Binomial	$n \cdot p$	$n \cdot p \cdot (1-p)$
Geometric	$\frac{1}{p}$	$\frac{1-p}{p^2}$

Exercise

Find the parameters of the Poisson, Normal, Uniform and exponential random variables.

Expectation of the function of a random variable

Expectation of the function of a random variable

Theorem

Expectation of the function of a random variable

Theorem

If X is a random variable with pmf $p()$,

Expectation of the function of a random variable

Theorem

If X is a random variable with pmf $p()$, and $g()$ is any real-valued function, then,

Expectation of the function of a random variable

Theorem

If X is a random variable with pmf $p()$, and $g()$ is any real-valued function, then,

$$E[g(X)] =$$

Expectation of the function of a random variable

Theorem

If X is a random variable with pmf $p()$, and $g()$ is any real-valued function, then,

$$E[g(X)] = \sum_{x:\ p(x)>0} g(x) \cdot p(x)$$

Expectation of the function of a random variable

Theorem

If X is a random variable with pmf $p()$, and $g()$ is any real-valued function, then,

$$E[g(X)] = \sum_{x:\ p(x)>0} g(x) \cdot p(x)$$

Joint Distributions

Joint Distributions

Joint distribution functions

Joint Distributions

Joint distribution functions

For any two random variables X and Y, the joint cumulative distribution function is defined as:

Joint Distributions

Joint distribution functions

For any two random variables X and Y, the joint cumulative distribution function is defined as:

$$F(a,b) = P(X \leq a, Y \leq b), \ -\infty < a, b < \infty$$

Joint Distributions

Joint distribution functions

For any two random variables X and Y, the joint cumulative distribution function is defined as:

$$F(a, b) = P(X \leq a, Y \leq b), \ -\infty < a, b < \infty$$

Joint Distributions

Joint distribution functions

For any two random variables X and Y, the joint cumulative distribution function is defined as:

$$F(a, b) = P(X \leq a, Y \leq b), \ -\infty < a, b < \infty$$

The distribution of X (or Y) can be obtained from the joint distribution as follows:

Joint Distributions

Joint distribution functions

For any two random variables X and Y, the joint cumulative distribution function is defined as:

$$F(a, b) = P(X \leq a, Y \leq b), \; -\infty < a, b < \infty$$

The distribution of X (or Y) can be obtained from the joint distribution as follows:

$$\begin{aligned} F_X(a) &= P(X \leq a) \\ &= P(X \leq a, Y \leq \infty) \\ &= F(a, \infty). \end{aligned}$$

Joint Distributions

Joint distribution functions

For any two random variables X and Y, the joint cumulative distribution function is defined as:

$$F(a, b) = P(X \leq a, Y \leq b), \ -\infty < a, b < \infty$$

The distribution of X (or Y) can be obtained from the joint distribution as follows:

$$\begin{aligned} F_X(a) &= P(X \leq a) \\ &= P(X \leq a, Y \leq \infty) \\ &= F(a, \infty). \end{aligned}$$

Note

In case X and Y are discrete random variables, we can define the joint probability mass function as:

Joint Distributions

Joint distribution functions

For any two random variables X and Y, the joint cumulative distribution function is defined as:

$$F(a, b) = P(X \leq a, Y \leq b), \ -\infty < a, b < \infty$$

The distribution of X (or Y) can be obtained from the joint distribution as follows:

$$\begin{aligned} F_X(a) &= P(X \leq a) \\ &= P(X \leq a, Y \leq \infty) \\ &= F(a, \infty). \end{aligned}$$

Note

In case X and Y are discrete random variables, we can define the joint probability mass function as:

$$p(x, y) = P(X = x, \ Y = y).$$

Independent Random Variables

Independent Random Variables

Definition

Two random variables X and Y are said to be independent, if

Independent Random Variables

Definition

Two random variables X and Y are said to be independent, if

$$F(a, b) = F_X(a) \cdot F_Y(b), \ \forall a, b.$$

Independent Random Variables

Definition

Two random variables X and Y are said to be independent, if

$$F(a,b) = F_X(a) \cdot F_Y(b), \ \forall a, b.$$

When X and Y are discrete, the above condition reduces to:

Independent Random Variables

Definition

Two random variables X and Y are said to be independent, if

$$F(a, b) = F_X(a) \cdot F_Y(b), \ \forall a, b.$$

When X and Y are discrete, the above condition reduces to:

$$p(x, y) = p_x(x) \cdot p_y(y)$$

Linearity of Expectation

Linearity of Expectation

Proposition

Linearity of Expectation

Proposition

Let $X_1, X_2, \ldots, X_n$ denote n random variables, defined over some probability space.

Linearity of Expectation

Proposition

Let $X_1, X_2, \ldots, X_n$ denote n random variables, defined over some probability space. Let $a_1, a_2, \ldots, a_n$ denote n constants. Then,

Linearity of Expectation

Proposition

Let $X_1, X_2, \ldots, X_n$ denote n random variables, defined over some probability space. Let $a_1, a_2, \ldots, a_n$ denote n constants. Then,

$$E[\sum_{i=1}^{n} a_i \cdot X_i] =$$

Linearity of Expectation

Proposition

Let $X_1, X_2, \ldots, X_n$ denote n random variables, defined over some probability space. Let $a_1, a_2, \ldots, a_n$ denote n constants. Then,

$$E[\sum_{i=1}^{n} a_i \cdot X_i] = \sum_{i=1}^{n} a_i \cdot E[X_i]$$

Linearity of Expectation

Proposition

Let $X_1, X_2, \ldots, X_n$ denote n random variables, defined over some probability space. Let $a_1, a_2, \ldots, a_n$ denote n constants. Then,

$$E[\sum_{i=1}^{n} a_i \cdot X_i] = \sum_{i=1}^{n} a_i \cdot E[X_i]$$

Note

Linearity of Expectation

Proposition

Let $X_1, X_2, \ldots, X_n$ denote n random variables, defined over some probability space. Let $a_1, a_2, \ldots, a_n$ denote n constants. Then,

$$E[\sum_{i=1}^{n} a_i \cdot X_i] = \sum_{i=1}^{n} a_i \cdot E[X_i]$$

Note

Note that linearity of expectation holds even when the random variables are **not** *independent.*

Linearity of Expectation

Proposition

Let $X_1, X_2, \ldots, X_n$ denote n random variables, defined over some probability space. Let $a_1, a_2, \ldots, a_n$ denote n constants. Then,

$$E[\sum_{i=1}^{n} a_i \cdot X_i] = \sum_{i=1}^{n} a_i \cdot E[X_i]$$

Note

Note that linearity of expectation holds even when the random variables are **not** *independent. For random variables X_1 and X_2, $Var(X_1 + X_2) = Var(X_1) + Var(X_2)$, only if X_1 and X_2 are independent.*

Linearity of Expectation

Proposition

Let $X_1, X_2, \ldots, X_n$ denote n random variables, defined over some probability space. Let $a_1, a_2, \ldots, a_n$ denote n constants. Then,

$$E[\sum_{i=1}^{n} a_i \cdot X_i] = \sum_{i=1}^{n} a_i \cdot E[X_i]$$

Note

Note that linearity of expectation holds even when the random variables are **not** *independent. For random variables X_1 and X_2, $Var(X_1 + X_2) = Var(X_1) + Var(X_2)$, only if X_1 and X_2 are independent. More generally,*

Linearity of Expectation

Proposition

Let $X_1,\ X_2,\ \ldots,\ X_n$ denote n random variables, defined over some probability space. Let $a_1,\ a_2,\ \ldots,\ a_n$ denote n constants. Then,

$$E[\sum_{i=1}^{n} a_i \cdot X_i] = \sum_{i=1}^{n} a_i \cdot E[X_i]$$

Note

Note that linearity of expectation holds even when the random variables are **not** *independent. For random variables X_1 and X_2, $Var(X_1 + X_2) = Var(X_1) + Var(X_2)$, only if X_1 and X_2 are independent. More generally,*

$$Var(X_1 + X_2) = Var(X_1) + Var(X_2) + 2 \cdot Cov(X_1, X_2).$$

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization
6. Financial Mathematics
   - Quantitative models
   - Problem Types

Concentration Inequalities

Concentration Inequalities

Tail bounds

Concentration Inequalities

Tail bounds

Consider the following problem:

Concentration Inequalities

Tail bounds

Consider the following problem: A fair coin is tossed n times. What is the probability that the number of heads is at least $\frac{3 \cdot n}{4}$?

Concentration Inequalities

Tail bounds

Consider the following problem: A fair coin is tossed n times. What is the probability that the number of heads is at least $\frac{3 \cdot n}{4}$? In general, the tail of a random X is the part of its pmf, that is away from its mean.

Concentration Inequalities

Tail bounds

Consider the following problem: A fair coin is tossed n times. What is the probability that the number of heads is at least $\frac{3 \cdot n}{4}$? In general, the tail of a random X is the part of its pmf, that is away from its mean.

Inequality	Known parameters	Tail bound
Markov	$X \geq 0, E[X]$	$P(X \geq a \cdot E[X]) \leq \frac{1}{a},\ a > 0$
Chebyshev	$E[X], Var(X)$	$P(\lvert X - E[X]\rvert \geq a \cdot E[X]) \leq \frac{Var(X)}{(a \cdot E[X])^2}, a > 0.$
Chernoff	X is binomial, $E[X]$	$P((X - E[X]) \geq \delta) \leq e^{-\frac{-2 \cdot \delta^2}{n}}, \delta > 0.$

Concentration Inequalities

Tail bounds

Consider the following problem: A fair coin is tossed n times. What is the probability that the number of heads is at least $\frac{3 \cdot n}{4}$? In general, the tail of a random X is the part of its pmf, that is away from its mean.

Inequality	Known parameters	Tail bound
Markov	$X \geq 0, E[X]$	$P(X \geq a \cdot E[X]) \leq \frac{1}{a}, \; a > 0$
Chebyshev	$E[X], Var(X)$	$P(\lvert X - E[X] \rvert \geq a \cdot E[X]) \leq \frac{Var(X)}{(a \cdot E[X])^2}, a > 0.$
Chernoff	X is binomial, $E[X]$	$P((X - E[X]) \geq \delta) \leq e^{-\frac{-2 \cdot \delta^2}{n}}, \delta > 0.$

Exercise

Find the tail bounds for the coin tossing problem using all three techniques.

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization
6. Financial Mathematics
   - Quantitative models
   - Problem Types

Optimization Theory

Optimization Theory

Fundamentals

Optimization Theory

Fundamentals

Given a function $f : \Re^n \rightarrow \Re$ and a set $S \subseteq \Re^n$, the problem of finding an $x^* \in \Re^n$ that solves

$$\begin{array}{rl} & \min_x f(x) \\ s.t. & x \in S \end{array}$$

is called an optimization problem.

Optimization Theory

Fundamentals

Given a function $f : \Re^n \rightarrow \Re$ and a set $S \subseteq \Re^n$, the problem of finding an $x^* \in \Re^n$ that solves

$$\begin{array}{rl} & \min_x f(x) \\ s.t. & x \in S \end{array}$$

is called an optimization problem.

Features of an optimization problem

Optimization Theory

Fundamentals

Given a function $f : \Re^n \rightarrow \Re$ and a set $S \subseteq \Re^n$, the problem of finding an $x^* \in \Re^n$ that solves

$$\begin{array}{rl} & \min_x f(x) \\ s.t. & x \in S \end{array}$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.

Optimization Theory

Fundamentals

Given a function $f : \Re^n \rightarrow \Re$ and a set $S \subseteq \Re^n$, the problem of finding an $x^* \in \Re^n$ that solves

$$\begin{array}{rc} & \min_x f(x) \\ s.t. & x \in S \end{array}$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.
- Objective function.

Optimization Theory

Fundamentals

Given a function $f : \Re^n \rightarrow \Re$ and a set $S \subseteq \Re^n$, the problem of finding an $x^* \in \Re^n$ that solves

$$\begin{array}{rc} & \min_x f(x) \\ s.t. & x \in S \end{array}$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.
- Objective function.
- Feasible region

Optimization Theory

Fundamentals

Given a function $f : \Re^n \rightarrow \Re$ and a set $S \subseteq \Re^n$, the problem of finding an $x^* \in \Re^n$ that solves

$$\begin{array}{rl} & \min_x f(x) \\ s.t. & x \in S \end{array}$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.
- Objective function.
- Feasible region (Infeasibility,

Optimization Theory

Fundamentals

Given a function $f : \Re^n \rightarrow \Re$ and a set $S \subseteq \Re^n$, the problem of finding an $x^* \in \Re^n$ that solves

$$\begin{array}{rl} & \min_x f(x) \\ s.t. & x \in S \end{array}$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.
- Objective function.
- Feasible region (Infeasibility, Unboundedness,

Optimization Theory

Fundamentals

Given a function $f : \Re^n \rightarrow \Re$ and a set $S \subseteq \Re^n$, the problem of finding an $x^* \in \Re^n$ that solves

$$\begin{array}{rl} & \min_x f(x) \\ s.t. & x \in S \end{array}$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.
- Objective function.
- Feasible region (Infeasibility, Unboundedness, Discrete).

Optimization Theory

Fundamentals

Given a function $f : \Re^n \rightarrow \Re$ and a set $S \subseteq \Re^n$, the problem of finding an $x^* \in \Re^n$ that solves

$$\begin{array}{rc} & \min_x f(x) \\ s.t. & x \in S \end{array}$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.
- Objective function.
- Feasible region (Infeasibility, Unboundedness, Discrete).
- Global minimizer (strict).

Optimization Theory

Fundamentals

Given a function $f : \Re^n \rightarrow \Re$ and a set $S \subseteq \Re^n$, the problem of finding an $x^* \in \Re^n$ that solves

$$\begin{array}{ll} & \min_x f(x) \\ s.t. & x \in S \end{array}$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.
- Objective function.
- Feasible region (Infeasibility, Unboundedness, Discrete).
- Global minimizer (strict).
- Local minimizer.

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization
6. Financial Mathematics
   - Quantitative models
   - Problem Types

Models of Optimization

Models of Optimization

Models

Models of Optimization

Models

1. Linear programming ($\min_{\mathbf{x}} \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x} \;\; \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \;\; \mathbf{x} \geq \mathbf{0}$).

Models of Optimization

Models

1. Linear programming ($\min_{\mathbf{x}} \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x} \;\; \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \;\; \mathbf{x} \geq \mathbf{0}$).
2. Non-linear programming ($\min_{\mathbf{x}} f(x) \;\; g_i(x) = 0, i \in \mathcal{E}, h_i(x) \geq 0, i \in \mathcal{I}$).

Models of Optimization

Models

1. Linear programming ($\min_{\mathbf{x}} \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x} \ \ \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}$).
2. Non-linear programming ($\min_{\mathbf{x}} f(x) \ \ g_i(x) = 0, i \in \mathcal{E}, h_i(x) \geq 0, i \in \mathcal{I}$).
3. Quadratic programming ($\min_{\mathbf{x}} \frac{1}{2}\mathbf{x}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{x} + \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x}$). Convexity, positive semidefinite matrices.

Models of Optimization

Models

1. Linear programming ($\min_{\mathbf{x}} \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x} \;\; \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \;\; \mathbf{x} \geq \mathbf{0}$).
2. Non-linear programming ($\min_{\mathbf{x}} f(x) \;\; g_i(x) = 0, i \in \mathcal{E}, h_i(x) \geq 0, i \in \mathcal{I}$).
3. Quadratic programming ($\min_{\mathbf{x}} \frac{1}{2}\mathbf{x}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{x} + \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x}$). Convexity, positive semidefinite matrices.
4. Conic optimization ($\mathbf{x} \in C$).

Models of Optimization

Models

1. Linear programming ($\min_{\mathbf{x}} \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x} \;\; \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \;\; \mathbf{x} \geq \mathbf{0}$).
2. Non-linear programming ($\min_{\mathbf{x}} f(x) \;\; g_i(x) = 0, i \in \mathcal{E}, h_i(x) \geq 0, i \in \mathcal{I}$).
3. Quadratic programming ($\min_{\mathbf{x}} \frac{1}{2}\mathbf{x}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{x} + \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x}$). Convexity, positive semidefinite matrices.
4. Conic optimization ($\mathbf{x} \in C$).
5. Integer programming ($\mathbf{x} \geq \mathbf{0}, \mathbf{x}$ integral). Binary programs.

Models of Optimization

Models

1. Linear programming ($\min_{\mathbf{x}} \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x} \;\; \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \;\; \mathbf{x} \geq \mathbf{0}$).
2. Non-linear programming ($\min_{\mathbf{x}} f(x) \;\; g_i(x) = 0, i \in \mathcal{E}, h_i(x) \geq 0, i \in \mathcal{I}$).
3. Quadratic programming ($\min_{\mathbf{x}} \frac{1}{2}\mathbf{x}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{x} + \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x}$). Convexity, positive semidefinite matrices.
4. Conic optimization ($\mathbf{x} \in C$).
5. Integer programming ($\mathbf{x} \geq \mathbf{0}, \mathbf{x}$ integral). Binary programs.
6. Dynamic programming.

Models of Optimization

Models

1. Linear programming ($\min_{\mathbf{x}} \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x} \;\; \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \;\; \mathbf{x} \geq \mathbf{0}$).
2. Non-linear programming ($\min_{\mathbf{x}} f(x) \;\; g_i(x) = 0, i \in \mathcal{E}, h_i(x) \geq 0, i \in \mathcal{I}$).
3. Quadratic programming ($\min_{\mathbf{x}} \frac{1}{2}\mathbf{x}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{x} + \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x}$). Convexity, positive semidefinite matrices.
4. Conic optimization ($\mathbf{x} \in C$).
5. Integer programming ($\mathbf{x} \geq \mathbf{0}, \mathbf{x}$ integral). Binary programs.
6. Dynamic programming.
7. Optimization with data uncertainty.

Models of Optimization

Models

1. Linear programming ($\min_{\mathbf{x}} \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x} \;\; \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \;\; \mathbf{x} \geq \mathbf{0}$).
2. Non-linear programming ($\min_{\mathbf{x}} f(x) \;\; g_i(x) = 0, i \in \mathcal{E}, h_i(x) \geq 0, i \in \mathcal{I}$).
3. Quadratic programming ($\min_{\mathbf{x}} \frac{1}{2}\mathbf{x}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{x} + \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x}$). Convexity, positive semidefinite matrices.
4. Conic optimization ($\mathbf{x} \in C$).
5. Integer programming ($\mathbf{x} \geq \mathbf{0}, \mathbf{x}$ integral). Binary programs.
6. Dynamic programming.
7. Optimization with data uncertainty.
    1. Stochastic programming.

Models of Optimization

Models

1. Linear programming ($\min_{\mathbf{x}} \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x} \;\; \mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \;\; \mathbf{x} \geq \mathbf{0}$).
2. Non-linear programming ($\min_{\mathbf{x}} f(x) \;\; g_i(x) = 0, i \in \mathcal{E}, h_i(x) \geq 0, i \in \mathcal{I}$).
3. Quadratic programming ($\min_{\mathbf{x}} \frac{1}{2}\mathbf{x}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{x} + \mathbf{c}^{\mathbf{T}} \cdot \mathbf{x}$). Convexity, positive semidefinite matrices.
4. Conic optimization ($\mathbf{x} \in C$).
5. Integer programming ($\mathbf{x} \geq \mathbf{0}, \mathbf{x}$ integral). Binary programs.
6. Dynamic programming.
7. Optimization with data uncertainty.
    1. Stochastic programming.
    2. Robust optimization.

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization
6. Financial Mathematics
   - Quantitative models
   - Problem Types

Financial Mathematics

Financial Mathematics

Principal issues

Financial Mathematics

Principal issues

1. Modern finance has become extremely technical.

Financial Mathematics

Principal issues

1. Modern finance has become extremely technical.
2. This field was originated by Markowitz (1950s) and Black, Schloes and Merton (1960s).

Outline

1. Linear Algebra
   - Vectors
   - Matrices
   - The Solution of Simultaneous Linear Equations
2. Convexity and Cones
   - Convexity
   - Cones
3. Probability and Expectation
   - Sample Space and Events
   - Defining Probabilities on Events
   - Conditional Probability
   - Random Variables
   - Concentration Inequalities
4. Basic optimization theory
   - Fundamentals
5. Models of Optimization
   - Tools of Optimization
6. Financial Mathematics
   - Quantitative models
   - Problem Types

Portfolio Selection and asset allocation

Portfolio Selection and asset allocation

Main Issues

Portfolio Selection and asset allocation

Main Issues

1. Select some from a number of securities.

Portfolio Selection and asset allocation

Main Issues

1. Select some from a number of securities.
2. Goal is to maximize return and minimize variance.

Portfolio Selection and asset allocation

Main Issues

1. Select some from a number of securities.
2. Goal is to maximize return and minimize variance.
3. Asset allocation.

Portfolio Selection and asset allocation

Main Issues

1. Select some from a number of securities.
2. Goal is to maximize return and minimize variance.
3. Asset allocation.
4. Index fund.

Portfolio Selection and asset allocation

Main Issues

1. Select some from a number of securities.
2. Goal is to maximize return and minimize variance.
3. Asset allocation.
4. Index fund.
5. Number of different models possible.

Pricing and hedging of options

Pricing and hedging of options

Main Issues

Pricing and hedging of options

Main Issues

1. Call/Put options.

Pricing and hedging of options

Main Issues

1. Call/Put options.
2. American/European style.

Pricing and hedging of options

Main Issues

1. Call/Put options.
2. American/European style.
3. How should an option be priced?

Pricing and hedging of options

Main Issues

1. Call/Put options.
2. American/European style.
3. How should an option be priced? Pricing problem.

Pricing and hedging of options

Main Issues

1. Call/Put options.
2. American/European style.
3. How should an option be priced? Pricing problem.
4. The replication approach.

Risk Management

Risk Management

Main Issues

Risk Management

Main Issues

1. Inherence of risk.

Risk Management

Main Issues

1. Inherence of risk.
2. Elimination versus management.

Risk Management

Main Issues

1. Inherence of risk.
2. Elimination versus management.
3. Quantitative measures and mathematical techniques.

Risk Management

Main Issues

1. Inherence of risk.
2. Elimination versus management.
3. Quantitative measures and mathematical techniques.
4. Some famous failures.

Risk Management

Main Issues

1. Inherence of risk.
2. Elimination versus management.
3. Quantitative measures and mathematical techniques.
4. Some famous failures.
5. Margin requirements.

Risk Management

Main Issues

1. Inherence of risk.
2. Elimination versus management.
3. Quantitative measures and mathematical techniques.
4. Some famous failures.
5. Margin requirements.
6. Typical problem - Optimize a performance measure,

Risk Management

Main Issues

1. Inherence of risk.
2. Elimination versus management.
3. Quantitative measures and mathematical techniques.
4. Some famous failures.
5. Margin requirements.
6. Typical problem - Optimize a performance measure, subject to the usual operating constraints,

Risk Management

Main Issues

1. Inherence of risk.
2. Elimination versus management.
3. Quantitative measures and mathematical techniques.
4. Some famous failures.
5. Margin requirements.
6. Typical problem - Optimize a performance measure, subject to the usual operating constraints, and the constraint that a particular risk measure does not exceed a threshold.

Asset/liability Management

Asset/liability Management

Main Issues

Asset/liability Management

Main Issues

1. Problems with the static approach.

Asset/liability Management

Main Issues

1. Problems with the static approach.
2. Should not penalize for above mean returns.

Asset/liability Management

Main Issues

1. Problems with the static approach.
2. Should not penalize for above mean returns.
3. Need for multi-period model.

Asset/liability Management

Main Issues

1. Problems with the static approach.
2. Should not penalize for above mean returns.
3. Need for multi-period model.
4. Optimization under uncertainty.

Asset/liability Management

Main Issues

1. Problems with the static approach.
2. Should not penalize for above mean returns.
3. Need for multi-period model.
4. Optimization under uncertainty.
5. Typical problem - What assets and in what quantities should the company hold in each period to maximize its wealth at the end of period T?