Linear Algebra

Department of Mathematics
Indian Institute of Technology Guwahati

January – May 2019

MA 102 (RA, RKS, MGPP, KVK)

The Vector Space $\mathbb{R}^n$

Topics:

- The vector space $\mathbb{R}^n$
- Inner product, length and angle

The Vector Space $\mathbb{R}^n$

Topics:

- The vector space $\mathbb{R}^n$
- Inner product, length and angle
- Linear combination
- Matrices and matrix-vector multiplication

The vector space $\mathbb{R}^n$

We define $\mathbb{R}^n$ to be the set of all ordered n-tuples of real numbers. Thus an n-tuple $\mathbf{v} \in \mathbb{R}^n$ is of the form

$$\text{row vector:} \quad \mathbf{v} = [v_1, \ldots, v_n] \ \text{ or } \ \text{column vector:} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

The vector space $\mathbb{R}^n$

We define $\mathbb{R}^n$ to be the set of all ordered n-tuples of real numbers. Thus an n-tuple $\mathbf{v} \in \mathbb{R}^n$ is of the form

$$\text{row vector:} \quad \mathbf{v} = [v_1, \ldots, v_n] \quad \text{or} \quad \text{column vector:} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

We always write an n-tuple in $\mathbb{R}^n$ as column vector. Thus

$$\mathbb{R}^n := \left\{ \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} : v_1, \ldots, v_n \in \mathbb{R} \right\}$$

The vector space $\mathbb{R}^n$

We define $\mathbb{R}^n$ to be the set of all ordered n-tuples of real numbers. Thus an n-tuple $\mathbf{v} \in \mathbb{R}^n$ is of the form

$$\text{row vector:} \quad \mathbf{v} = [v_1, \ldots, v_n] \quad \text{or} \quad \text{column vector:} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

We always write an n-tuple in $\mathbb{R}^n$ as column vector. Thus

$$\mathbb{R}^n := \left\{ \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} : v_1, \ldots, v_n \in \mathbb{R} \right\}$$

Transpose: $[v_1, \ldots, v_n]^\top = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ and $\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}^\top = [v_1, \ldots, v_n]$.

Algebraic properties of $\mathbb{R}^n$

Define addition and scalar multiplication on $\mathbb{R}^n$ as follows:

$$\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix} \text{ and } \alpha \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} \alpha u_1 \\ \vdots \\ \alpha u_n \end{bmatrix} \text{ for } \alpha \in \mathbb{R}.$$

Algebraic properties of $\mathbb{R}^n$

Define addition and scalar multiplication on $\mathbb{R}^n$ as follows:

$$\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix} \text{ and } \alpha \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} \alpha u_1 \\ \vdots \\ \alpha u_n \end{bmatrix} \text{ for } \alpha \in \mathbb{R}.$$

Then for $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $\mathbb{R}^n$ and scalars α, β in $\mathbb{R}$, the following hold:

1. Commutativity: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
2. Associativity: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
3. Identity: $\mathbf{u} + \mathbf{0} = \mathbf{u}$
4. Inverse: $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$

Algebraic properties of $\mathbb{R}^n$

Define addition and scalar multiplication on $\mathbb{R}^n$ as follows:

$$\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix} \text{ and } \alpha \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} \alpha u_1 \\ \vdots \\ \alpha u_n \end{bmatrix} \text{ for } \alpha \in \mathbb{R}.$$

Then for $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $\mathbb{R}^n$ and scalars α, β in $\mathbb{R}$, the following hold:

1. Commutativity: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
2. Associativity: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
3. Identity: $\mathbf{u} + \mathbf{0} = \mathbf{u}$
4. Inverse: $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$
5. Distributivity : $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$
6. Distributivity : $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}$
7. Associativity: $\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u}$
8. Identity: $1\mathbf{u} = \mathbf{u}$.

The vector space $\mathbb{R}^n$

The set $\mathbb{R}^n$ equipped with vector addition "$+$" and scalar multiplication "$\cdot$" is called a vector space.

Exercise: Let $\mathbf{u}, \mathbf{v}, \mathbf{x}$ be vectors in $\mathbb{R}^3$.
(a) Simplify $3\mathbf{u} + (5\mathbf{v} - 2\mathbf{u}) + 2(\mathbf{u} - \mathbf{v})$
(b) Solve $5\mathbf{x} - \mathbf{u} = 2(\mathbf{u} + 2\mathbf{x})$ for $\mathbf{x}$.

The vector space $\mathbb{R}^n$

The set $\mathbb{R}^n$ equipped with vector addition "$+$" and scalar multiplication "$\cdot$" is called a vector space.

Exercise: Let $\mathbf{u}, \mathbf{v}, \mathbf{x}$ be vectors in $\mathbb{R}^3$.
(a) Simplify $3\mathbf{u} + (5\mathbf{v} - 2\mathbf{u}) + 2(\mathbf{u} - \mathbf{v})$
(b) Solve $5\mathbf{x} - \mathbf{u} = 2(\mathbf{u} + 2\mathbf{x})$ for $\mathbf{x}$.

Length and angle: Length, distance and angle can all be described by using the notion of inner product (dot product) of two vectors.

Definition: If $\mathbf{u} := [u_1, \ldots, u_n]^\top$ and $\mathbf{v} := [v_1, \ldots, v_n]^\top$ are vectors in $\mathbb{R}^n$ then the inner product $\langle \mathbf{u}, \mathbf{v} \rangle$ is defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle := u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

The inner product $\langle \mathbf{u}, \mathbf{v} \rangle$ is also written as the dot product $\mathbf{u} \bullet \mathbf{v}$.

The vector space $\mathbb{R}^n$

The set $\mathbb{R}^n$ equipped with vector addition "$+$" and scalar multiplication "$\cdot$" is called a vector space.

Exercise: Let $\mathbf{u}, \mathbf{v}, \mathbf{x}$ be vectors in $\mathbb{R}^3$.
(a) Simplify $3\mathbf{u} + (5\mathbf{v} - 2\mathbf{u}) + 2(\mathbf{u} - \mathbf{v})$
(b) Solve $5\mathbf{x} - \mathbf{u} = 2(\mathbf{u} + 2\mathbf{x})$ for $\mathbf{x}$.

Length and angle: Length, distance and angle can all be described by using the notion of inner product (dot product) of two vectors.

Definition: If $\mathbf{u} := [u_1, \ldots, u_n]^\top$ and $\mathbf{v} := [v_1, \ldots, v_n]^\top$ are vectors in $\mathbb{R}^n$ then the inner product $\langle \mathbf{u}, \mathbf{v} \rangle$ is defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle := u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

The inner product $\langle \mathbf{u}, \mathbf{v} \rangle$ is also written as the dot product $\mathbf{u} \bullet \mathbf{v}$.

Example: If $\mathbf{u} := [1, 2, -3]^\top$ and $\mathbf{v} := [-3, 5, 2]^\top$ then

$$\langle \mathbf{u}, \mathbf{v} \rangle = 1 \cdot (-3) + 2 \cdot 5 + (-3) \cdot 2 = 1.$$

Length and angle

Theorem: Let $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$ and let $\alpha \in \mathbb{R}$. Then

1. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ and $\langle \mathbf{u}, \mathbf{u} \rangle = 0 \iff \mathbf{u} = \mathbf{0}$.
2. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
3. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$
4. $\langle \alpha \mathbf{u}, \mathbf{v} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle$.

Length and angle

Theorem: Let $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$ and let $\alpha \in \mathbb{R}$. Then

1. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ and $\langle \mathbf{u}, \mathbf{u} \rangle = 0 \iff \mathbf{u} = \mathbf{0}$.
2. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
3. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$
4. $\langle \alpha \mathbf{u}, \mathbf{v} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle$.

Definition: The norm (or length) of a vector $\mathbf{v} := [v_1, \ldots, v_n]^\top$ in $\mathbb{R}^n$ is a nonnegative number $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| := \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{v_1^2 + \cdots + v_n^2}.$$

Length and angle

Theorem: Let $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$ and let $\alpha \in \mathbb{R}$. Then

1. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ and $\langle \mathbf{u}, \mathbf{u} \rangle = 0 \iff \mathbf{u} = \mathbf{0}$.
2. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
3. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$
4. $\langle \alpha \mathbf{u}, \mathbf{v} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle$.

Definition: The norm (or length) of a vector $\mathbf{v} := [v_1, \ldots, v_n]^\top$ in $\mathbb{R}^n$ is a nonnegative number $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| := \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{v_1^2 + \cdots + v_n^2}.$$

Cauchy-Schwarz Inequality: Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$. Then

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|.$$

Proof: Define $p(t) := \langle \mathbf{u} + t\mathbf{v}, \mathbf{u} + t\mathbf{v} \rangle$ for $t \in \mathbb{R}$. Then $p(t) = \|u\|^2 + 2t\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2 t^2 \geq 0$ for all $t \in \mathbb{R}$ yields the result.

Length and angle

Theorem: Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$ and let $\alpha \in \mathbb{R}$. Then

1. Positive definite: $\|\mathbf{u}\| = 0 \iff \mathbf{u} = \mathbf{0}$
2. Positive homogeneity: $\|\alpha \mathbf{u}\| = |\alpha| \, \|\mathbf{u}\|$
3. Triangle inequality: $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Length and angle

Theorem: Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$ and let $\alpha \in \mathbb{R}$. Then

1. Positive definite: $\|\mathbf{u}\| = 0 \iff \mathbf{u} = \mathbf{0}$
2. Positive homogeneity: $\|\alpha \mathbf{u}\| = |\alpha| \, \|\mathbf{u}\|$
3. Triangle inequality: $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Unit vector: A vector $\mathbf{v}$ in $\mathbb{R}^n$ is called a unit vector if $\|\mathbf{v}\| = 1$. If $\mathbf{u}$ is a nonzero vector then $\mathbf{v} := \frac{1}{\|\mathbf{u}\|}\mathbf{u}$ is a unit vector in the direction of $\mathbf{u}$. Here $\mathbf{v}$ is referred to as normalization of $\mathbf{u}$.

Example: The vectors $\mathbf{e}_1 := [1, 0, 0]^\top, \mathbf{e}_2 := [0, 1, 0]^\top$ and $\mathbf{e}_3 := [0, 0, 1]^\top$ are unit vectors in $\mathbb{R}^3$. These vectors are called standard unit vectors.

Length and angle

Theorem: Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$ and let $\alpha \in \mathbb{R}$. Then

1. Positive definite: $\|\mathbf{u}\| = 0 \iff \mathbf{u} = \mathbf{0}$
2. Positive homogeneity: $\|\alpha\mathbf{u}\| = |\alpha|\,\|\mathbf{u}\|$
3. Triangle inequality: $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Unit vector: A vector $\mathbf{v}$ in $\mathbb{R}^n$ is called a unit vector if $\|\mathbf{v}\| = 1$. If $\mathbf{u}$ is a nonzero vector then $\mathbf{v} := \frac{1}{\|\mathbf{u}\|}\mathbf{u}$ is a unit vector in the direction of $\mathbf{u}$. Here $\mathbf{v}$ is referred to as normalization of $\mathbf{u}$.

Example: The vectors $\mathbf{e}_1 := [1,0,0]^\top, \mathbf{e}_2 := [0,1,0]^\top$ and $\mathbf{e}_3 := [0,0,1]^\top$ are unit vectors in $\mathbb{R}^3$. These vectors are called standard unit vectors.

Distance: The distance $\mathrm{d}(\mathbf{u},\mathbf{v})$ between two vectors $\mathbf{u} := [u_1,\ldots,u_n]^\top$ and $\mathbf{v} := [v_1,\ldots,v_n]^\top$ in $\mathbb{R}^n$ is defined by

$$\mathrm{d}(\mathbf{u},\mathbf{v}) := \|\mathbf{u}-\mathbf{v}\| = \sqrt{(u_1-v_1)^2 + \cdots + (u_n-v_n)^2}.$$

Angle and orthogonality

Let $\mathbf{u}$ and $\mathbf{v}$ be nonzero vectors in $\mathbb{R}^n$. Consider the triangle with sides $\mathbf{u}, \mathbf{v}$ and $\mathbf{u}-\mathbf{v}$. Let θ be the angle between $\mathbf{u}$ and $\mathbf{v}$. Then the low of cosines applied to the triangle yields

$$\|\mathbf{u}-\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta.$$

Angle and orthogonality

Let $\mathbf{u}$ and $\mathbf{v}$ be nonzero vectors in $\mathbb{R}^n$. Consider the triangle with sides $\mathbf{u}, \mathbf{v}$ and $\mathbf{u} - \mathbf{v}$. Let θ be the angle between $\mathbf{u}$ and $\mathbf{v}$. Then the low of cosines applied to the triangle yields

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\| \, \|\mathbf{v}\| \cos\theta.$$

Expanding $\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 - 2\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2$ gives us

$$\langle \mathbf{u}, \mathbf{v}\rangle = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos\theta \Longrightarrow \cos\theta = \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\| \, \|\mathbf{v}\|}.$$

Angle and orthogonality

Let $\mathbf{u}$ and $\mathbf{v}$ be nonzero vectors in $\mathbb{R}^n$. Consider the triangle with sides $\mathbf{u}, \mathbf{v}$ and $\mathbf{u} - \mathbf{v}$. Let θ be the angle between $\mathbf{u}$ and $\mathbf{v}$. Then the low of cosines applied to the triangle yields

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\| \, \|\mathbf{v}\| \cos\theta.$$

Expanding $\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 - 2\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2$ gives us

$$\langle \mathbf{u}, \mathbf{v}\rangle = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos\theta \Longrightarrow \cos\theta = \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\| \, \|\mathbf{v}\|}.$$

Definition: Two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are said to be orthogonal to each other if $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.

Angle and orthogonality

Let $\mathbf{u}$ and $\mathbf{v}$ be nonzero vectors in $\mathbb{R}^n$. Consider the triangle with sides $\mathbf{u}, \mathbf{v}$ and $\mathbf{u} - \mathbf{v}$. Let θ be the angle between $\mathbf{u}$ and $\mathbf{v}$. Then the low of cosines applied to the triangle yields

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta.$$

Expanding $\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 - 2\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2$ gives us

$$\langle \mathbf{u}, \mathbf{v}\rangle = \|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta \Longrightarrow \cos\theta = \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|}.$$

Definition: Two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are said to be orthogonal to each other if $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.

Example: The vectors $\mathbf{u} := [1, 1, -2]^\top$ and $\mathbf{v} := [3, 1, 2]^\top$ in $\mathbb{R}^3$ are orthogonal as $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.

Angle and orthogonality

Let $\mathbf{u}$ and $\mathbf{v}$ be nonzero vectors in $\mathbb{R}^n$. Consider the triangle with sides $\mathbf{u}, \mathbf{v}$ and $\mathbf{u} - \mathbf{v}$. Let θ be the angle between $\mathbf{u}$ and $\mathbf{v}$. Then the low of cosines applied to the triangle yields

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\| \, \|\mathbf{v}\| \cos\theta.$$

Expanding $\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 - 2\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2$ gives us

$$\langle \mathbf{u}, \mathbf{v}\rangle = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos\theta \Longrightarrow \cos\theta = \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\| \, \|\mathbf{v}\|}.$$

Definition: Two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are said to be orthogonal to each other if $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.

Example: The vectors $\mathbf{u} := [1, 1, -2]^\top$ and $\mathbf{v} := [3, 1, 2]^\top$ in $\mathbb{R}^3$ are orthogonal as $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.

Pythagoras' Theorem: Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$. Then

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \iff \langle \mathbf{u}, \mathbf{v}\rangle = 0.$$

Matrices

Definition: A matrix is an array of numbers called entries or elements of the matrix. The size of a matrix A is a description of the number of rows and columns of the matrix A. An $m \times n$ matrix A has m rows and n columns and is of the form

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

Matrices

Definition: A matrix is an array of numbers called entries or elements of the matrix. The size of a matrix A is a description of the number of rows and columns of the matrix A. An $m \times n$ matrix A has m rows and n columns and is of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Let $\mathbf{a}_j := [a_{1j}, \ldots, a_{mj}]^\top$ be the j-th column of A for $j = 1 : n$. Then we represent A as $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$.

Matrices

Definition: A matrix is an array of numbers called entries or elements of the matrix. The size of a matrix A is a description of the number of rows and columns of the matrix A. An $m \times n$ matrix A has m rows and n columns and is of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Let $\mathbf{a}_j := [a_{1j}, \ldots, a_{mj}]^\top$ be the j-th column of A for $j = 1 : n$.
Then we represent A as $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$.
Let $\mathbf{A}_i := [a_{i1}, a_{i2}, \ldots, a_{in}]$ be the i-th row of A for $i = 1 : m$. Then we represent A as $A = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_m \end{bmatrix}$.

Linear combination

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$. Let α and β be scalars. Adding $\alpha\mathbf{u}$ and $\beta\mathbf{v}$ gives the linear combination $\alpha\mathbf{u} + \beta\mathbf{v}$.

Example: Let $\mathbf{u} := [1, 1, -1]^\top, \mathbf{v} := [2, 3, 4]^\top$ and $\mathbf{w} := [4, 5, 2]^\top$. Then $\mathbf{w} = 2\mathbf{u} + \mathbf{v}$. Thus $\mathbf{w}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

Linear combination

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$. Let α and β be scalars. Adding $\alpha\mathbf{u}$ and $\beta\mathbf{v}$ gives the linear combination $\alpha\mathbf{u} + \beta\mathbf{v}$.

Example: Let $\mathbf{u} := [1, 1, -1]^\top, \mathbf{v} := [2, 3, 4]^\top$ and $\mathbf{w} := [4, 5, 2]^\top$. Then $\mathbf{w} = 2\mathbf{u} + \mathbf{v}$. Thus $\mathbf{w}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

Definition: Let $\mathbf{v}_1, \dots, \mathbf{v}_m$ be vectors in $\mathbb{R}^n$ and let $\alpha_1, \dots, \alpha_m$ be scalars. Then the vector $\mathbf{u} := \alpha_1\mathbf{v}_1 + \cdots + \alpha_m\mathbf{v}_m$ is called a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_m$.

Linear combination

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$. Let α and β be scalars. Adding $\alpha\mathbf{u}$ and $\beta\mathbf{v}$ gives the linear combination $\alpha\mathbf{u} + \beta\mathbf{v}$.

Example: Let $\mathbf{u} := [1, 1, -1]^\top, \mathbf{v} := [2, 3, 4]^\top$ and $\mathbf{w} := [4, 5, 2]^\top$. Then $\mathbf{w} = 2\mathbf{u} + \mathbf{v}$. Thus $\mathbf{w}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

Definition: Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be vectors in $\mathbb{R}^n$ and let $\alpha_1, \ldots, \alpha_m$ be scalars. Then the vector $\mathbf{u} := \alpha_1\mathbf{v}_1 + \cdots + \alpha_m\mathbf{v}_m$ is called a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$.

Problem: Let $\mathbf{a}_1, \ldots, \mathbf{a}_n$ and $\mathbf{b}$ be vectors in $\mathbb{R}^m$. Find scalars $x_1, \ldots, x_n$, if exist, such that $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$.

Example: Vector equation

$$x_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}.$$

Matrix times vector

We rewrite the linear combination $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ using a matrix. Set $A := \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$ and $\mathbf{x} := [x_1, \ldots, x_n]^\top$. We define the matrix A times the vector $\mathbf{x}$ to be the same as the combination $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$.

Matrix times vector

We rewrite the linear combination $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ using a matrix. Set $A := \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$ and $\mathbf{x} := [x_1, \ldots, x_n]^\top$. We define the matrix A times the vector $\mathbf{x}$ to be the same as the combination $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$.

Definition: Matrix-vector multiplication

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n.$$

The matrix A acts on the vector $\mathbf{x}$ and the result $A\mathbf{x}$ is a linear combination of the columns of A.

Matrix times vector

We rewrite the linear combination $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ using a matrix. Set $A := \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$ and $\mathbf{x} := [x_1, \ldots, x_n]^\top$. We define the matrix A times the vector $\mathbf{x}$ to be the same as the combination $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$.

Definition: Matrix-vector multiplication

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n.$$

The matrix A acts on the vector $\mathbf{x}$ and the result $A\mathbf{x}$ is a linear combination of the columns of A.

Example: Compact notation for vector equation

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}.$$

Matrix times vector

A row vector $\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix}$ is a $1 \times n$ matrix. Therefore

$$\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = a_{i1}x_1 + \cdots + a_{in}x_n.$$

Matrix times vector

A row vector $\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix}$ is a $1 \times n$ matrix. Therefore

$$\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = a_{i1}x_1 + \cdots + a_{in}x_n.$$

Example: Matrix-vector multiplication in two ways

$$\begin{aligned} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= x_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} x_1 \\ x_1 + x_2 \\ x_2 + x_2 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \mathbf{x} \\ \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \mathbf{x} \\ \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \mathbf{x} \end{bmatrix} \end{aligned}$$

Matrix-vector multiplication

More generally

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ldots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \end{bmatrix} \mathbf{x} \\ \vdots \\ \begin{bmatrix} a_{m1} & \cdots & a_{mn} \end{bmatrix} \mathbf{x} \end{bmatrix}.$$

Matrix-vector multiplication

More generally

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ldots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \end{bmatrix} \mathbf{x} \\ \vdots \\ \begin{bmatrix} a_{m1} & \cdots & a_{mn} \end{bmatrix} \mathbf{x} \end{bmatrix}.$$

Now represent $A := \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$ by its rows: $A = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_m \end{bmatrix}$.

Then we have

$$A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1\mathbf{x} \\ \vdots \\ \mathbf{A}_m\mathbf{x} \end{bmatrix}.$$

Matrices and Vectors in Data Mining

Problem: Given a few key words, retrieve relevant information from a large database.

Matrices and Vectors in Data Mining

Problem: Given a few key words, retrieve relevant information from a large database.

Term-Document Matrix: Term-document matrices are used in information retrieval. Consider the five documents.

Matrices and Vectors in Data Mining

Problem: Given a few key words, retrieve relevant information from a large database.

Term-Document Matrix: Term-document matrices are used in information retrieval. Consider the five documents.

Doc. 1:	The Google matrix G is a model of the Internet.
Doc. 2:	G_{ij} is nonzero if there is a link from web page j to i.
Doc. 3:	The Google matrix G is used to rank all web pages.
Doc. 4:	The ranking is done by solving a matrix eigenvalue problem.
Doc. 5:	England dropped out of the top 10 in the FIFA ranking.

Matrices and Vectors in Data Mining

Problem: Given a few key words, retrieve relevant information from a large database.

Term-Document Matrix: Term-document matrices are used in information retrieval. Consider the five documents.

Doc. 1:	The Google matrix G is a model of the Internet.
Doc. 2:	G_{ij} is nonzero if there is a link from web page j to i.
Doc. 3:	The Google matrix G is used to rank all web pages.
Doc. 4:	The ranking is done by solving a matrix eigenvalue problem.
Doc. 5:	England dropped out of the top 10 in the FIFA ranking.

The key words or terms are colored blue. The set of terms is called a Dictionary. Counting the frequency of terms in each document, we obtain a term-document matrix.

Term-document matrix

Term	Doc. 1	Doc. 2	Doc. 3	Doc. 4	Doc. 5
eigenvalue	0	0	0	1	0
England	0	0	0	0	1
FIFA	0	0	0	0	1
Google	1	0	1	0	0
Internet	1	0	0	0	0
link	0	1	0	0	0
matrix	1	0	1	1	0
page	0	1	1	0	0
rank	0	0	1	1	1
web	0	1	1	0	1

Term-document matrix

Term	Doc. 1	Doc. 2	Doc. 3	Doc. 4	Doc. 5
eigenvalue	0	0	0	1	0
England	0	0	0	0	1
FIFA	0	0	0	0	1
Google	1	0	1	0	0
Internet	1	0	0	0	0
link	0	1	0	0	0
matrix	1	0	1	1	0
page	0	1	1	0	0
rank	0	0	1	1	1
web	0	1	1	0	1

Each document is represented by a 10×5 column of the term-document matrix A which is a vector in $\mathbb{R}^{10}$.

Query vector

Suppose that we want to find all documents that are relevant to the query ranking of web pages. This is represented by a query vector, constructed in the way as the term-document matrix, using the same dictionary:

$$\mathbf{v} := \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}^\top \in \mathbb{R}^{10}.$$

Query vector

Suppose that we want to find all documents that are relevant to the query ranking of web pages. This is represented by a query vector, constructed in the way as the term-document matrix, using the same dictionary:

$$\mathbf{v} := \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}^\top \in \mathbb{R}^{10}.$$

Thus the query itself is a document. The information retrieval task can now be formulated as a mathematical problem.

Query vector

Suppose that we want to find all documents that are relevant to the query ranking of web pages. This is represented by a query vector, constructed in the way as the term-document matrix, using the same dictionary:

$$\mathbf{v} := \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}^\top \in \mathbb{R}^{10}.$$

Thus the query itself is a document. The information retrieval task can now be formulated as a mathematical problem.

> Problem: Find the columns of A that are close (in some sense) to the query vector $\mathbf{v}$.

Query matching (use of dot product)

Query matching is the process of finding all documents that are relevant to a particular query $\mathbf{v}$. The cosine of angle between two vectors is often used to determine relevant documents:

$$\cos\theta_j := \frac{\langle Ae_j, \mathbf{v}\rangle}{\|\mathbf{v}\|\,\|Ae_j\|} > \mathrm{tol}$$

where Ae_j is the j-th column of A and tol is a predefined tolerance. Thus $\cos\theta_j > \mathrm{tol} \Rightarrow Ae_j$ is relevant.

Query matching (use of dot product)

Query matching is the process of finding all documents that are relevant to a particular query $\mathbf{v}$. The cosine of angle between two vectors is often used to determine relevant documents:

$$\cos\theta_j := \frac{\langle Ae_j, \mathbf{v}\rangle}{\|\mathbf{v}\| \, \|Ae_j\|} > \mathrm{tol}$$

where Ae_j is the j-th column of A and tol is a predefined tolerance. Thus $\cos\theta_j > \mathrm{tol} \Rightarrow Ae_j$ is relevant.

Consider the term-document matrix A and the query ("ranking of web pages") vector $\mathbf{v}$. Then the cosines measures of the query and the original data are given by

$$[0, 0.6667, 0.7746, 0.3333, 0.3333]^T$$

which shows that Doc. 2-3 are most relevant.

***