Chapter 3: Vector Spaces

March 16, 2009

Lecture 16

1 Vector spaces

A **vector space** is a non-empty set V of objects, called **vectors**, on which are defined two operations, called **addition** and **scalar multiplication**: for any vectors $\boldsymbol{u}, \boldsymbol{v}$ in V, the sum $\boldsymbol{u}+\boldsymbol{v}$ is in V; for a vector $\boldsymbol{u}$ in V and a scalar c (real number), the scalar multiple $c\boldsymbol{u}$ is in V; subject to the following axioms listed below. The axioms must hold for all vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ in V and scalars c and d.

1. $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$,
2. $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$,
3. There is a vector $\mathbf{0}$ such that $\boldsymbol{u}+\mathbf{0}=\boldsymbol{u}$,
4. For any vector $\boldsymbol{u}$ there is a vector $-\boldsymbol{u}$ such that $\boldsymbol{u}+(-\boldsymbol{u})=\mathbf{0}$;
5. $c(\boldsymbol{u}+\boldsymbol{v})=c\boldsymbol{u}+c\boldsymbol{v}$,
6. $(c+d)\boldsymbol{u}=c\boldsymbol{u}+d\boldsymbol{u}$,
7. $c(d\boldsymbol{u})=(cd)\boldsymbol{u}$,
8. $1\boldsymbol{u}=\boldsymbol{u}$.

The vector $\mathbf{0}$ is called the **zero vector** of V. The vector $-\boldsymbol{u}$ is called the **negative vector** of $\boldsymbol{u}$.

By definition of vector space it is easy to see that for any vector $\boldsymbol{u}$ and scalar c,

$$0\boldsymbol{u}=\mathbf{0}, \quad c\mathbf{0}=\mathbf{0}, \quad -\boldsymbol{u}=(-1)\boldsymbol{u}.$$

For instance,

$$\begin{aligned}
0\boldsymbol{u} &\overset{(3)}{=} 0\boldsymbol{u}+\mathbf{0} \overset{(4)}{=} 0\boldsymbol{u}+(0\boldsymbol{u}+(-0\boldsymbol{u})) \overset{(2)}{=} (0\boldsymbol{u}+0\boldsymbol{u})+(-0\boldsymbol{u}) \\
&\overset{(6)}{=} (0+0)\boldsymbol{u}+(-0\boldsymbol{u}) = 0\boldsymbol{u}+(-0\boldsymbol{u}) \overset{(4)}{=} \mathbf{0}; \\
c\mathbf{0} &= c(0\boldsymbol{u}) \overset{(7)}{=} (c0)\boldsymbol{u} = 0\boldsymbol{u} = \mathbf{0}; \\
-\boldsymbol{u} &= -\boldsymbol{u}+\mathbf{0} = -\boldsymbol{u}+(1-1)\boldsymbol{u} = -\boldsymbol{u}+\boldsymbol{u}+(-1)\boldsymbol{u} = \mathbf{0}+(-1)\boldsymbol{u} = (-1)\boldsymbol{u}.
\end{aligned}$$

Example 1.1. (a) The Euclidean space $\mathbb{R}^n$ is a vector space under the ordinary addition and scalar multiplication.

(b) The set $\mathbf{P}_n$ of all polynomials of degree less than or equal to n is a vector space under the ordinary addition and scalar multiplication of polynomials.

(c) The set $\mathbf{M}(m, n)$ of all $m \times n$ matrices is a vector space under the ordinary addition and scalar multiplication of matrices.

(d) The set $C[a, b]$ of all continuous functions on the closed interval $[a, b]$ is a vector space under the ordinary addition and scalar multiplication of functions. (We do not study this kind of spaces here.)

Definition 1.1. Let V and W be vector spaces and $W \subseteq V$. If the addition and scalar multiplication in W are the same as the addition and scalar multiplication in V, then W is called a **subspace** of V.

If H is a subspace of V, then H is closed for the addition and scalar multiplication of V: if $\boldsymbol{u}, \boldsymbol{v} \in H$ and scalar $c \in \mathbb{R}$, then

$$\boldsymbol{u} + \boldsymbol{v} \in H, \quad c\boldsymbol{v} \in H.$$

Theorem 1.2. *Let H be a non-empty subset of a vector space V. Then H is a subspace of V if and only if H is closed under addition and scalar multiplication, i.e.,*

(a) *For any vectors $\boldsymbol{u}, \boldsymbol{v} \in H$, we have $\boldsymbol{u} + \boldsymbol{v} \in H$,*

(b) *For any scalar c and a vector $\boldsymbol{v} \in H$, we have $c\boldsymbol{v} \in H$.*

Example 1.2. (a) For a vector space V, the set $\{\mathbf{0}\}$ of the zero vector is a subspace, called the **zero subspace** of V. The whole space V is a subspace of V. The subspaces $\{\mathbf{0}\}$ and V are called the **trivial subspaces** of V.

(b) For an $m \times n$ matrix A, the set of solutions of the linear system $A\boldsymbol{x} = \mathbf{0}$ is a subspace of $\mathbb{R}^n$. However, if $\boldsymbol{b} \neq \mathbf{0}$, the set of solutions of the system $A\boldsymbol{x} = \boldsymbol{b}$ is *not* a subspace of $\mathbb{R}^n$.

(c) For any linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, the image

$$T(\mathbb{R}^n) = \{T(\boldsymbol{x}) \mid \boldsymbol{x} \in \mathbb{R}^n\}$$

of T is a subspace of $\mathbb{R}^m$, and the inverse image

$$T^{-1}(\mathbf{0}) = \{\boldsymbol{x} \in \mathbb{R}^n \mid T(\boldsymbol{x}) = \mathbf{0}\}$$

is a subspace of $\mathbb{R}^n$.

(d) The set $\mathbf{R}^n = \{(a_1, \ldots, a_n) \mid a_1, \ldots, a_n \in \mathbb{R}\}$ is a vector space under the addition and scalar multiplication:

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n),$$
$$c(a_1, \ldots, a_n) = (ca_1, \ldots, ca_n).$$

2 Subspace spanned by a set

Let V be a vector space. Given vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_p$ in V, we can form a new vector

$$c_1\boldsymbol{v}_1 + \cdots + c_p\boldsymbol{v}_p,$$

called the **linear combination** of $\boldsymbol{v}_1, \dots, \boldsymbol{v}_p$ with scalars $c_1, \dots, c_p$ (real numbers). The **span** of vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_p$ in V is the set

$$\mathrm{Span}\,\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_p\} := \big\{c_1\boldsymbol{v}_1 + \cdots + c_p\boldsymbol{v}_p \mid c_1, \dots, c_p \in \mathbb{R}\big\}.$$

Example 2.1. Given two vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ in a vector space V, let $H = \mathrm{Span}\,\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$. Show that H is a subspace.

Theorem 2.1. *If $\boldsymbol{v}_1, \dots, \boldsymbol{v}_p$ are vectors in a vector space V, then* $\mathrm{Span}\,\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_p\}$ *is a subspace of V, called the* **subspace spanned** *by $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_p\}$.*

Let H be a subspace of a vector space V. A **spanning set** for H is a set $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_p\}$ of vectors in H such that

$$H = \mathrm{Span}\,\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_p\}.$$

3 Null spaces and column spaces of matrices

Lecture 17

Definition 3.1. Let A be an $m \times n$ matrix. The **null space** of A, denoted by $\mathrm{Nul}\,A$, is the space of solutions of the linear system $A\boldsymbol{x} = \boldsymbol{0}$, i.e.,

$$\mathrm{Nul}\,A := \{\boldsymbol{x} \in \mathbb{R}^n \mid A\boldsymbol{x} = \boldsymbol{0}\}.$$

The **column space** of A, denoted by $\mathrm{Col}\,A$, is the span of the column vectors of A, i.e., writing $A = [\boldsymbol{a}_1, \dots, \boldsymbol{a}_n]$,

$$\mathrm{Col}\,A := \mathrm{Span}\,\{\boldsymbol{a}_1, \dots, \boldsymbol{a}_n\}.$$

Let $T : \mathbb{R}^n \to \mathbb{R}^m$, $T(\boldsymbol{x}) = A\boldsymbol{x}$, be a linear transformation. Then $\mathrm{Nul}\,A$ is the set of inverse images of $\boldsymbol{0}$ under T and $\mathrm{Col}\,A$ is the image of T, that is,

$$\mathrm{Nul}\,A = T^{-1}(\boldsymbol{0}) \quad \text{and} \quad \mathrm{Col}\,A = T(\mathbb{R}^n).$$

4 Linear transformations

Definition 4.1. A **linear transformation** T from a vector space V into a vector space W is a rule that assigns to each vector $\boldsymbol{x}$ in V a unique vector $T(\boldsymbol{x})$ in W, such that

(a) $T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$ for all $\boldsymbol{u}, \boldsymbol{v}$ in V, and

(b) $T(c\boldsymbol{u}) = cT(\boldsymbol{u})$ for all $\boldsymbol{u}$ in V and all scalars c.

The **kernel** (or **null space**) of T is the set of all vectors $\boldsymbol{u}$ such that $T(\boldsymbol{u}) = \boldsymbol{0}$. The **range** of T is the set of all vectors in W of the form $T(\boldsymbol{x})$ for some $\boldsymbol{x}$ in V.

Example 4.1. (a) Let A is an $m \times m$ matrix and B an $n \times n$ matrix. The function

$$F : \mathbf{M}(m,n) \to \mathbf{M}(m,n), \quad F(X) = AXB$$

is a linear transformation. For instance, for $m = n = 2$, let

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}.$$

Then $F : \mathbf{M}(2,2) \to \mathbf{M}(2,2)$ is given by

$$F(X) = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 2x_1 + 2x_2 + 4x_3 + 4x_4 & x_1 + 3x_2 + 2x_3 + 6x_4 \\ 2x_1 + 2x_2 + 6x_3 + 6x_4 & x_1 + 3x_2 + 3x_3 + 9x_4 \end{bmatrix}.$$

(b) The function $D : \mathbf{P}_3 \to \mathbf{P}_2$, defined by $D(p(t)) = \frac{d}{dt}p(t)$, i.e.,

$$D\big(a_0 + a_1 t + a_2 t^2 + a_3 t^3\big) = a_1 + 2a_2 t + 3a_3 t^2,$$

is a linear transformation.

5 Independent sets and bases

Definition 5.1. Given a vector space V. Vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ in V are said to be **linearly independent** provided that, if there are scalars $c_1, \ldots, c_p$ such that

$$c_1\boldsymbol{v}_1 + \cdots + c_p\boldsymbol{v}_p = \mathbf{0},$$

then $c_1 = \cdots = c_p = 0$. The vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ are said to be **linearly dependent** if there are some scalars $c_1, \ldots, c_p$, not all zero, such that

$$c_1\boldsymbol{v}_1 + \cdots + c_p\boldsymbol{v}_p = \mathbf{0}.$$

Any family of vectors that contain the zero vector $\mathbf{0}$ is linearly dependent. A single vector $\boldsymbol{v}$ is linearly independent if and only if $\boldsymbol{v} \neq \mathbf{0}$.

Theorem 5.2. *Let $T : V \to W$ be a linear transformation. Given vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ in V.*

(a) *If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ are linearly dependent, then $T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_p)$ are linearly dependent;*

(b) *If $T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_p)$ are linearly independent, then $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ are linearly independent.*

Theorem 5.3. *Vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ $(p \geq 2)$ are linearly dependent if and only if at least one of the vectors is a linear combination of the others, i.e., there is a vector $\boldsymbol{v}_k$ such that*

$$\boldsymbol{v}_k = a_1\boldsymbol{v}_1 + \cdots + a_{i-1}\boldsymbol{v}_{k-1} + a_{k+1}\boldsymbol{v}_{i+1} + \cdots + a_p\boldsymbol{v}_p.$$

Moreover, if $\boldsymbol{v}_1 \neq \mathbf{0}$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ are linearly dependent, then at least one vector $\boldsymbol{v}_k$ $(k \geq 2)$ is a linear combination of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}$.

Proof. Since the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ are linearly dependent, there are constants $c_1, \ldots, c_p$, not all zero, such that

$$c_1\boldsymbol{v} + \cdots + c_k\boldsymbol{v}_k + \cdots + c_p\boldsymbol{v}_p = \mathbf{0}.$$

Since $c_1, \ldots, c_p$ are not all zero, say, $c_k \neq 0$. Then

$$\boldsymbol{v}_k = \left(-\frac{c_1}{c_k}\right)\boldsymbol{v}_1 + \cdots + \left(-\frac{c_{k-1}}{c_k}\right)\boldsymbol{v}_{k-1} + \left(-\frac{c_{k+1}}{c_k}\right)\boldsymbol{v}_{k+1} + \cdots + \left(-\frac{c_p}{c_i}\right)\boldsymbol{v}_p.$$

□

Example 5.1. Let $p_1(t) = 2$, $p_2(t) = t + 1$, $p_3(t) = 3t + 1$. Then $\{p_1(t), p_2(t), p_3(t)\}$ is linearly dependent. $[p_3(t) = 3p_2(t) - p_1(t).]$

Theorem 5.4. *Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ be a subset of independent vectors in a vector space V. If a vector $\boldsymbol{v}$ can be written in two linear combinations of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$, say,*

$$\boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_p\boldsymbol{v}_p = d_1\boldsymbol{v}_1 + \cdots + d_p\boldsymbol{v}_p,$$

then $c_1 = d_1, \ldots, c_p = d_p$.

Proof. If $(c_1, \ldots, c_p) \neq (d_1, \ldots, d_p)$, then one of the entries in $(c_1 - d_1, \ldots, c_p - d_p)$ is non-zero, and

$$(c_1 - d_1)\boldsymbol{v}_1 + \cdots + (c_p - d_p)\boldsymbol{v}_p = \mathbf{0}.$$

This means that the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ are linearly dependent. This is a contradiction. □

Definition 5.5. Let H be a subspace of a vector space V. An ordered set $\mathcal{B} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ of vectors in V is called a **basis** for H if

(a) $\mathcal{B}$ is a linearly independent set, and

(b) $\mathcal{B}$ spans H, i.e., $H = \operatorname{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$.

Example 5.2. (a) The set $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ is a basis of $\mathbb{R}^n$, called the **standard basis** of $\mathbb{R}^n$.

(b) The set $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is basis of $\mathbb{R}^3$, where

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \quad \boldsymbol{v}_3 = \begin{bmatrix} 3 \\ 4 \\ 6 \end{bmatrix}.$$

Example 5.3. (a) The set $\{1, t, t^2\}$ is basis of $\mathbf{P}_2$.

(b) The set $\{1, t + 1, t^2 + t\}$ is basis of $\mathbf{P}_2$.

(c) The set $\{1, t + 1, t - 1\}$ is not a basis of $\mathbf{P}_2$.

Example 5.4. The vector space $\mathbf{M}(2, 2)$ of 2×2 matrices has a basis

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

The following set

$$\mathcal{C} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$$

is also a basis for $\mathbf{M}(2, 2)$.

Proposition 5.6 (Spanning Theorem)**.** *Given vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ in a vector space V; let $H = \operatorname{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$.*

(a) *If one of the given vectors, say, $\boldsymbol{v}_k$, is a linear combination of the other vectors, then*

$$H = \operatorname{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_p\}.$$

(b) *If $H \neq \{\mathbf{0}\}$, then some subset S of vectors is a basis for H.*

Proof. It is clear that Span $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_p\}$ is contained in H. Write

$$\boldsymbol{v}_k = c_1\boldsymbol{v}_1 + \cdots + c_{k-1}\boldsymbol{v}_{k-1} + c_{k+1}\boldsymbol{v}_{k+1} + \cdots + c_p\boldsymbol{v}_p.$$

Then for any vector $\boldsymbol{v} = a_1\boldsymbol{v}_1 + \cdots + a_k\boldsymbol{v}_k + \cdots + a_p\boldsymbol{v}_p$ in H, we have

$$\begin{aligned} \boldsymbol{v} &= (a_1 + a_kc_1)\boldsymbol{v}_1 + \cdots + (a_{k-1} + a_kc_{k-1})\boldsymbol{v}_{k-1} \\ &\quad +(a_{k+1} + a_kc_{k+1})\boldsymbol{v}_{k+1} + \cdots + (a_p + a_kc_p)\boldsymbol{v}_p. \end{aligned}$$

This means that $\boldsymbol{v}$ is contained in Span $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_p\}$. □

6 Bases for $\operatorname{Nul} A$ and $\operatorname{Col} A$

Example 6.1. Consider the matrix

$$A = \begin{bmatrix} 1 & 4 & 0 & 2 & 4 \\ -1 & -4 & 1 & -3 & -2 \\ 2 & 8 & 1 & 3 & 10 \\ 1 & 4 & 1 & 1 & 6 \end{bmatrix} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4, \boldsymbol{a}_5].$$

Its reduced row echelon form is the matrix

$$B = \begin{bmatrix} 1 & 4 & 0 & 2 & 4 \\ 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = [\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3, \boldsymbol{b}_4, \boldsymbol{b}_5].$$

Since $A\boldsymbol{x} = \boldsymbol{0}$ is equivalent to $B\boldsymbol{x} = \boldsymbol{0}$, that is,

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3 + x_4\boldsymbol{a}_4 + x_5\boldsymbol{a}_5 = \boldsymbol{0} \quad \Longleftrightarrow \quad x_1\boldsymbol{b}_1 + x_2\boldsymbol{b}_2 + x_3\boldsymbol{b}_3 + x_4\boldsymbol{b}_4 + x_5\boldsymbol{b}_5 = \boldsymbol{0}.$$

This means that the linear relations among the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4, \boldsymbol{a}_5$ are the same as the linear relations among the vectors $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3, \boldsymbol{b}_4, \boldsymbol{b}_5$. For instance,

$$\begin{array}{lcl} \boldsymbol{b}_2 = 4\boldsymbol{b}_1 & \longleftrightarrow & \boldsymbol{a}_2 = 4\boldsymbol{a}_1 \\ \boldsymbol{b}_4 = 2\boldsymbol{b}_1 - \boldsymbol{b}_3 & \longleftrightarrow & \boldsymbol{a}_4 = 2\boldsymbol{a}_1 - \boldsymbol{a}_3 \\ \boldsymbol{b}_5 = 4\boldsymbol{b}_1 + 2\boldsymbol{b}_3 & \longleftrightarrow & \boldsymbol{a}_4 = 4\boldsymbol{a}_1 + 2\boldsymbol{a}_3. \end{array}$$

This shows that row operations *do not* change the linear relations among the column vectors of a matrix.

Note 1. Let A and B be matrix such that $A \sim B$, i.e., A is equivalent to B. Then

$$\operatorname{Nul} A = \operatorname{Nul} B, \quad \operatorname{Row} A = \operatorname{Row} B, \quad \text{but} \quad \operatorname{Col} A \neq \operatorname{Col} B.$$

Theorem 6.1 (Column Space Theorem)**.** *The column vectors of a matrix A corresponding to its pivot positions form a basis of* $\operatorname{Col} A$.

Proof. Let $B = [\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n]$ denote the reduced row echelon form of $A = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n]$. Let $\boldsymbol{b}_{i_1}, \boldsymbol{b}_{i_2}, \ldots, \boldsymbol{b}_{i_k}$ be the column vectors of B containing the pivot positions. It is clear that $\boldsymbol{b}_{i_1}, \boldsymbol{b}_{i_2}, \ldots, \boldsymbol{b}_{i_k}$ are linearly independent and every column vector of B is a linear combination of the vectors $\boldsymbol{b}_{i_1}, \boldsymbol{b}_{i_2}, \ldots, \boldsymbol{b}_{i_k}$.

Let $\boldsymbol{a}_{i_1}, \boldsymbol{a}_{i_2}, \ldots, \boldsymbol{a}_{i_k}$ be the corresponding column vectors of A. It suffices to prove that a linear relation for $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n$ is also a linear relation for $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$, and vice versa. Notice that a linear relation among the vectors $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n$ is just a solution of the system $B\boldsymbol{x} = \boldsymbol{0}$; and the systems $A\boldsymbol{x} = \boldsymbol{0}$ and $B\boldsymbol{x} = \boldsymbol{0}$ have the same solution set. Thus $\boldsymbol{a}_{i_1}, \boldsymbol{a}_{i_2}, \ldots, \boldsymbol{a}_{i_k}$ are linearly independent and every column vector of A is a linear combination of $\boldsymbol{a}_{i_1}, \boldsymbol{a}_{i_2}, \ldots, \boldsymbol{a}_{i_k}$. So $\boldsymbol{a}_{i_1}, \boldsymbol{a}_{i_2}, \ldots, \boldsymbol{a}_{i_k}$ form a basis of $\operatorname{Col} A$. □

7 Coordinate systems

Lecture 18

Theorem 7.1. *Let $\mathcal{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ be a basis of a vector space V. Then for each vector $\boldsymbol{v}$ in V, there exists a unique set of scalars $c_1, \ldots, c_n$ such that*

$$\boldsymbol{v} = c_1\boldsymbol{b}_1 + \cdots + c_n\boldsymbol{b}_n.$$

Proof. Trivial. □

Definition 7.2. Let $\mathcal{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ be a basis of a vector space V. The **coordinates of $\boldsymbol{v}$ relative to the basis** $\mathcal{B}$ (or **$\mathcal{B}$-coordinates of $\boldsymbol{v}$**) are scalars $c_1, \ldots, c_n$ such that

$$\boldsymbol{v} = c_1\boldsymbol{b}_1 + \cdots + c_n\boldsymbol{b}_n.$$

The vector

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

is called the **coordinate vector of $\boldsymbol{v}$ relative to $\mathcal{B}$** (or **$\mathcal{B}$-coordinate vector of $\boldsymbol{v}$**).

Example 7.1. Any two linearly independent vectors of $\mathbb{R}^2$ form a basis for $\mathbb{R}^2$. For instance, the set

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

is basis of $\mathbb{R}^2$. The vector $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ has the $\mathcal{B}$-coordinate vector $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$. However, the coordinate vector of $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ is just itself under the standard basis

$$\mathcal{E} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}.$$

Theorem 7.3. *Let $\mathcal{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ be a basis of a subspace H of $\mathbb{R}^n$. Let $P_{\mathcal{B}}$ be the matrix*

$$P_{\mathcal{B}} = [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n].$$

Then for any vector $\boldsymbol{v}$ in $\mathbb{R}^n$,

$$\boldsymbol{v} = P_{\mathcal{B}}[\boldsymbol{v}]_{\mathcal{B}}.$$

The matrix $P_{\mathcal{B}}$, which transfers the $\mathcal{B}$-coordinate vector $[\boldsymbol{v}]_{\mathcal{B}}$ of $\boldsymbol{v}$ to its standard coordinate vector $\boldsymbol{v} = [\boldsymbol{v}]_{\mathcal{E}}$, is called the **change-of-coordinate matrix from $\mathcal{B}$ to $\mathcal{E}$**.

Proof. Let $\boldsymbol{v} = c_1\boldsymbol{b}_1 + \cdots + c_n\boldsymbol{b}_n$. Then

$$\boldsymbol{v} = [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n] \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = P_{\mathcal{B}}\,[\boldsymbol{v}]_{\mathcal{B}}.$$

□

Theorem 7.4. *Let $\mathcal{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ be a basis of a vector space V. Then the coordinate transformation,*

$$V \to \mathbb{R}^n, \quad \boldsymbol{v} \mapsto [\boldsymbol{v}]_{\mathcal{B}},$$

is linear, one-to-one, and onto.

Proof. For vectors $\boldsymbol{v}, \boldsymbol{w}$ of V and scalar a, if

$$\boldsymbol{v} = c_1\boldsymbol{b}_1 + \cdots + c_n\boldsymbol{b}_n,$$

$$\boldsymbol{w} = d_1\boldsymbol{b}_1 + \cdots + d_n\boldsymbol{b}_n,$$

then

$$\boldsymbol{v} + \boldsymbol{w} = (c_1 + d_1)\boldsymbol{b}_1 + \cdots + (c_n + d_n)\boldsymbol{b}_n,$$
$$a\boldsymbol{v} = a(c_1\boldsymbol{b}_1 + \cdots + c_n\boldsymbol{b}_n) = (ac_1)\boldsymbol{b}_1 + \cdots + (ac_n)\boldsymbol{b}_n.$$

Thus

$$[\boldsymbol{v} + \boldsymbol{w}]_{\mathcal{B}} = \begin{bmatrix} c_1 + d_1 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix},$$

$$[c\boldsymbol{v}]_{\mathcal{B}} = \begin{bmatrix} cc_1 \\ \vdots \\ cc_n \end{bmatrix} = c\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = c\,[\boldsymbol{v}]_{\mathcal{B}}.$$

So the coordinate transformation is a linear transformation.

Now for any vector

$$\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \quad \text{in} \quad \mathbb{R}^n,$$

consider the vector $\boldsymbol{v} = c_1\boldsymbol{b}_1 + \cdots + c_n\boldsymbol{b}_n$ in V. The coordinate vector of $\boldsymbol{v}$ relative to $\mathcal{B}$ is

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Thus the transformation is onto. The linear independence of $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ implies that the transformation is also one-to-one. □

A one-to-one and onto linear transformation from a vector space V to a vector space W is called an **isomorphism**.

Example 7.2. The vector space $\mathbf{P}_3$ of polynomials of degree at most 3 in variable t is isomorphic to the vector space $\mathbb{R}^4$, and $\{1, t, t^2, t^3\}$ is a basis of $\mathbf{P}_3$.

Proof. The map $F : \mathbf{P}_3 \to \mathbb{R}^4$, defined by

$$F[p(t)] = F\left(c_0 + c_1 t + c_2 t^2 + c_3 t^3\right) = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix},$$

is a one-to-one linear transformation from $\mathbf{P}_3$ onto $\mathbb{R}^4$. □

Example 7.3. The vector space $\mathbf{M}(2,2)$ of 2×2 matrices is isomorphic to $\mathbb{R}^4$, and the matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

form a basis of $\mathbf{M}(2,2)$. In fact, the map $F : \mathbf{M}(2,2) \to \mathbb{R}^4$, defined by

$$F(M) = F\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix},$$

is a one-to-one linear transformation from $\mathbf{M}(2,2)$ onto $\mathbb{R}^4$.

Theorem 7.5. *Let $\mathcal{B} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ be a basis of a vector space V. Then any set of V consisting more than n vectors are linearly dependent.*

Proof. Let $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p\}$ be a set of vectors with $p > n$. Since any set of more than n vectors of $\mathbb{R}^n$ is linearly dependent, the vectors $[\boldsymbol{u}_1]_{\mathcal{B}}, \ldots, [\boldsymbol{u}_p]_{\mathcal{B}}$ of $\mathbb{R}^n$ must be linearly dependent. Then there exist constants $c_1, \ldots, c_p$, not all zero, such that

$$c_1 [\boldsymbol{u}_1]_{\mathcal{B}} + \cdots + c_p [\boldsymbol{u}_p]_{\mathcal{B}} = \mathbf{0}.$$

Thus

$$[c_1\boldsymbol{u}_1 + \cdots + c_p\boldsymbol{u}_p]_{\mathcal{B}} = c_1 [\boldsymbol{u}_1]_{\mathcal{B}} + \cdots + c_p [\boldsymbol{u}_p]_{\mathcal{B}} = \mathbf{0} = [\mathbf{0}]_{\mathcal{B}}.$$

Note that the coordinate transformation is one-to-one. It follows that

$$c_1\boldsymbol{u}_1 + \cdots + c_p\boldsymbol{u}_p = \mathbf{0}.$$

This means that the vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ are linearly dependent by definition. □

Theorem 7.6. *If $\mathcal{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ and $\mathcal{C} = \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_p\}$ are bases of a vector space V, then $n = p$.*

Proof. Suppose $n < p$. By Theorem 7.5, $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_p\}$ is linearly dependent, contrary to the properties for a basis. Thus $n \geq p$. A similar argument shows that $n \leq p$. Hence $n = p$. □

8 Dimensions of vector spaces

Lecture 19

A vector space V is said to be **finite-dimensional** if it can be spanned by a set of finite number of vectors. The dimension of V, denoted by $\dim V$, is the number of vectors of a basis of V. The dimension of the zero vector space $\{\mathbf{0}\}$ is zero. If V cannot be spanned by any finite set of vectors, then V is said to be **infinite-dimensional**.

Theorem 8.1. *Let H be a subspace of a finite-dimensional vector space V. Then any linearly independent subset of H can be expanded to a basis of H. Moreover, H is finite dimensional and $\dim H \leq \dim V$.*

Proof. Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ be a set of linearly independent vectors of H. If $\operatorname{Span} S = H$, then S is a basis of H by definition. Otherwise, there exists a vector $\boldsymbol{v}_{p+1}$ in H such that $\boldsymbol{v}_{p+1}$ is not in $\operatorname{Span} S$. Then $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p, \boldsymbol{v}_{p+1}\}$ is a linearly independent set of H. Now set

$$S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p, \boldsymbol{v}_{p+1}\}.$$

If $\operatorname{Span} S = H$, then S is a basis of H. Otherwise, continue to add one vector of $H - \operatorname{Span} S$ to S in this way until $\operatorname{Span} S = H$. Since H is of finite dimensional, the extension ends in finite number of steps. □

Theorem 8.2 (Basis Theorem)**.** *Given a set $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ of n vectors of an n-dimensional vector space V.*

(a) *If $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is linearly independent, then $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is a basis of V.*

(b) *If $\operatorname{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\} = V$, then $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is a basis of V.*

Proof. (a) By Theorem 8.1, S can be extended to a basis of V. Since S has n vectors and all bases have the same number of vectors. It follows that no vectors were added to S to be extended a basis of V. Hence S itself must be a basis.

(b) We need to show that S is linearly independent. Note that if S is not a basis, then S is linearly dependent. Thus S contains a linearly independent proper subset S' such that $\operatorname{Span} S' = V$. So S' is a basis of V; therefore $\#(S') \geq n$, contradict to $\#(S') < n$. □

9 Rank

Theorem 9.1. *For any rectangular matrix A,*

$$\dim \operatorname{Row} A = \dim \operatorname{Col} A = \#\{\text{pivot positions of } A\}.$$

Definition 9.2. The **rank** of a rectangular matrix A is the number pivot positions of A, that is, the dimension of the row space and the column space of A.

The **rank** of a linear transformation $T : V \to W$ is the dimension of the subspace $T(V)$.

Theorem 9.3 (Rank Theorem)**.** *For any $m \times n$ matrix A,*

$$\operatorname{rank} A + \dim \operatorname{Nul} A = n.$$

Proof. The rank of A is the number of pivot positions of A and the dimension of the null space of A is the number of free variables of the system $A\boldsymbol{x} = \boldsymbol{0}$. It is clear that

$$\#\{\text{pivot positions}\} + \#\{\text{free variables}\} = n.$$

□

Theorem 9.4. *(a) If two matrices A and B are row equivalent, then* $\operatorname{Row} A = \operatorname{Row} B$.

(b) If B is in echelon form, then the non-zero rows of B form a basis of $\operatorname{Row} B$.

Theorem 9.5. *Let A be an $n \times n$ invertible matrix. Then*

$$\dim \operatorname{Row} A = \dim \operatorname{Col} A = \operatorname{rank} A = n,$$

$$\dim \operatorname{Nul} A = 0.$$

Proof. The invertibility of A implies that the number of pivot positions of A is n. So $\operatorname{rank} A = n$ and $\dim \operatorname{Nul} A = 0$. □

10 Matrices of linear transformations

Week 10

Definition 10.1. Let V be an n-dimensional vector space V with basis $\mathcal{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n\}$. Let $T : V \to V$ be a linear transformation of V to itself. Write

$$\begin{cases} T(\boldsymbol{b}_1) &= a_{11}\boldsymbol{b}_1 + a_{21}\boldsymbol{b}_2 + \cdots + a_{n1}\boldsymbol{b}_n \\ T(\boldsymbol{b}_2) &= a_{12}\boldsymbol{b}_1 + a_{22}\boldsymbol{b}_2 + \cdots + a_{n2}\boldsymbol{b}_n \\ &\vdots \\ T(\boldsymbol{b}_n) &= a_{1n}\boldsymbol{b}_1 + a_{2n}\boldsymbol{b}_2 + \cdots + a_{nn}\boldsymbol{b}_n \end{cases}$$

The $n \times n$ matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \Big[[T(\boldsymbol{b}_1)]_{\mathcal{B}},\ [T(\boldsymbol{b}_2)]_{\mathcal{B}},\ \dots,\ [T(\boldsymbol{b}_n)]_{\mathcal{B}}\Big].$$

is called the **matrix of T relative to the basis $\mathcal{B}$.**

Let $T : V \to V$ be a linear transformation with matrix A relative to a basis $\mathcal{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n\}$. For each vector $\boldsymbol{v}$ of V, let

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad [T(\boldsymbol{v})]_{\mathcal{B}} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

be the coordinate vectors of $\boldsymbol{v}$ and $T(\boldsymbol{v})$ respectively, i.e.,

$$\boldsymbol{v} = x_1\boldsymbol{b}_1 + x_2\boldsymbol{b}_2 + \cdots + x_n\boldsymbol{b}_n = [\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

$$T(\boldsymbol{v}) = y_1\boldsymbol{b}_1 + y_2\boldsymbol{b}_2 + \cdots + y_n\boldsymbol{b}_n = [\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Then

$$T(\boldsymbol{v}) = x_1T(\boldsymbol{b}_1) + x_2T(\boldsymbol{b}_2) + \cdots + x_nT(\boldsymbol{b}_n) = [T(\boldsymbol{b}_1), T(\boldsymbol{b}_2), \dots, T(\boldsymbol{b}_n)] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Note that

$$[T(\boldsymbol{b}_1), T(\boldsymbol{b}_2), \dots, T(\boldsymbol{b}_n)] = [\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

It follows that

$$T(\boldsymbol{v}) = [\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Comparing the above formulas, we have

$$[T(\boldsymbol{v})]_{\mathcal{B}} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Example 10.1. Let $T : \mathbf{P}_3 \to \mathbf{P}_3$ be defined by $T(p(t)) = \frac{\mathrm{d}}{\mathrm{d}t}p(t)$. It is easy to check that T $(= \frac{\mathrm{d}}{\mathrm{d}t})$ is a linear transformation. It is clear that $\mathbf{P}_3$ has a basis $\mathcal{B} = \{1,\ t,\ t^2,\ t^3\}$, and

$$\begin{cases} T(1) &= 0 &= 0\cdot 1 + 0\cdot t + 0\cdot t^2 + 0\cdot t^3 \\ T(t) &= 1 &= 1\cdot 1 + 0\cdot t + 0\cdot t^2 + 0\cdot t^3 \\ T(t^2) &= 2t &= 0\cdot 1 + 2\cdot t + 0\cdot t^2 + 0\cdot t^3 \\ T(t^3) &= 3t^2 &= 0\cdot 1 + 0\cdot t + 3\cdot t^2 + 0\cdot t^3 \end{cases}$$

The matrix of T relative to the basis $\mathcal{B}$ is the matrix

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Example 10.2. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A\boldsymbol{x}.$$

The set $\mathcal{P} = \{\boldsymbol{p}_1, \boldsymbol{p}_2\} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ is a basis of $\mathbb{R}^2$. What is the matrix of T relative to the basis $\mathcal{P}$? By calculation we see that

$$T(\boldsymbol{p}_1) = 3\boldsymbol{p}_1, \quad T(\boldsymbol{p}_2) = \boldsymbol{p}_2.$$

So the matrix of T relative to the basis $\mathcal{P}$ is the matrix $B = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$. Let $P = [\boldsymbol{p}_1, \boldsymbol{p}_2]$. Then

$$\begin{aligned} AP &= A[\boldsymbol{p}_1, \boldsymbol{p}_2] = [A\boldsymbol{p}_1, A\boldsymbol{p}_2] = [3\boldsymbol{p}_1, \boldsymbol{p}_2] \\ &= \begin{bmatrix} 3 & 1 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} = PB. \end{aligned}$$

This means that

$$B = P^{-1}AP.$$

Definition 10.2. An $n \times n$ matrix A is said to be **similar** to an $n \times n$ matrix B if there is an invertible matrix P such that

$$P^{-1}AP = B.$$

Theorem 10.3. *For a finite dimensional vector space V and a linear transformation $T : V \to V$, the matrices of T relative to various bases are similar. In other words, the matrices of the same linear transformation from a vector space to itself under different bases are similar.*