2

Linear Independence, Span, and Bases

2.1 Span and Linear Independence **2**-1
2.2 Basis and Dimension of a Vector Space **2**-3
2.3 Direct Sum Decompositions **2**-4
2.4 Matrix Range, Null Space, Rank, and the Dimension Theorem **2**-6
2.5 Nonsingularity Characterizations **2**-9
2.6 Coordinates and Change of Basis **2**-10
2.7 Idempotence and Nilpotence **2**-12
References **2**-12

Mark Mills
Central College

2.1 Span and Linear Independence

Let V be a vector space over a field F.

Definitions:

A **linear combination** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in V$ is a sum of scalar multiples of these vectors; that is, $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$, for some scalar coefficients $c_1, c_2, \ldots, c_k \in F$. If S is a set of vectors in V, a linear combination of vectors in S is a vector of the form $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$ with $k \in \mathbb{N}, \mathbf{v}_i \in S, c_i \in F$. Note that S may be finite or infinite, but a linear combination is, by definition, a finite sum. The zero vector is defined to be a linear combination of the empty set.

When all the scalar coefficients in a linear combination are 0, it is a **trivial linear combination.** A sum over the empty set is also a trivial linear combination.

The **span** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in V$ is the set of all linear combinations of these vectors, denoted by $\text{Span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$. If S is a (finite or infinite) set of vectors in V, then the span of S, denoted by $\text{Span}(S)$, is the set of all linear combinations of vectors in S.

If $V = \text{Span}(S)$, then S **spans** the vector space V.

A (finite or infinite) set of vectors S in V is **linearly independent** if the only linear combination of distinct vectors in S that produces the zero vector is a trivial linear combination. That is, if $\mathbf{v}_i$ are distinct vectors in S and $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$, then $c_1 = c_2 = \cdots = c_k = 0$. Vectors that are not linearly independent are **linearly dependent**. That is, there exist distinct vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in S$ and $c_1, c_2, \ldots, c_k$ not all 0 such that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$.

Facts: The following facts can be found in [Lay03, Sections 4.1 and 4.3].

1. $\text{Span}(\emptyset) = \{\mathbf{0}\}$.
2. A linear combination of a single vector $\mathbf{v}$ is simply a scalar multiple of $\mathbf{v}$.
3. In a vector space V, $\text{Span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ is a subspace of V.
4. Suppose the set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ spans the vector space V. If one of the vectors, say $\mathbf{v}_i$, is a linear combination of the remaining vectors, then the set formed from S by removing $\mathbf{v}_i$ still spans V.
5. Any single nonzero vector is linearly independent.
6. Two nonzero vectors are linearly independent if and only if neither is a scalar multiple of the other.
7. If S spans V and $S \subseteq T$, then T spans V.
8. If T is a linearly independent subset of V and $S \subseteq T$, then S is linearly independent.
9. Vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly dependent if and only if $\mathbf{v}_i = c_1\mathbf{v}_1 + \cdots + c_{i-1}\mathbf{v}_{i-1} + c_{i+1}\mathbf{v}_{i+1} + \cdots + c_k\mathbf{v}_k$, for some $1 \leq i \leq k$ and some scalars $c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_k$. A set S of vectors in V is linearly dependent if and only if there exists $\mathbf{v} \in S$ such that $\mathbf{v}$ is a linear combination of other vectors in S.
10. Any set of vectors that includes the zero vector is linearly dependent.

Examples:

1. Linear combinations of $\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix} \in \mathbb{R}^2$ are vectors of the form $c_1\begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_2\begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} c_1 \\ -c_1 + 3c_2 \end{bmatrix}$, for any scalars $c_1, c_2 \in \mathbb{R}$. Any vector of this form is in $\text{Span}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix}\right)$. In fact, $\text{Span}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix}\right) = \mathbb{R}^2$ and these vectors are linearly independent.
2. If $\mathbf{v} \in \mathbb{R}^n$ and $\mathbf{v} \neq \mathbf{0}$, then geometrically $\text{Span}(\mathbf{v})$ is a line in $\mathbb{R}^n$ through the origin.
3. Suppose $n \geq 2$ and $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$ are linearly independent vectors. Then geometrically $\text{Span}(\mathbf{v}_1, \mathbf{v}_2)$ is a plane in $\mathbb{R}^n$ through the origin.
4. Any polynomial $p(x) \in \mathbb{R}[x]$ of degree less than or equal to 2 can easily be seen to be a linear combination of $1, x$, and x^2. However, $p(x)$ is also a linear combination of $1, 1+x$, and $1+x^2$. So $\text{Span}(1, x, x^2) = \text{Span}(1, 1+x, 1+x^2) = \mathbb{R}[x; 2]$.
5. The n vectors $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$ span F^n, for any field F. These vectors are also linearly independent.
6. In $\mathbb{R}^2$, $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$ are linearly independent. However, $\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ 5 \end{bmatrix}$ are linearly dependent, because $\begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 3 \end{bmatrix}$.
7. The infinite set $\{1, x, x^2, \ldots, x^n, \ldots\}$ is linearly independent in $F[x]$, for any field F.
8. In the vector space of continuous real-valued functions on the real line, $\mathcal{C}(\mathbb{R})$, the set $\{\sin(x), \sin(2x), \ldots, \sin(nx), \cos(x), \cos(2x), \ldots, \cos(nx)\}$ is linearly independent for any $n \in \mathbb{N}$. The infinite set $\{\sin(x), \sin(2x), \ldots, \sin(nx), \ldots, \cos(x), \cos(2x), \ldots, \cos(nx), \ldots\}$ is also linearly independent in $\mathcal{C}(\mathbb{R})$.

Applications:

1. The homogeneous differential equation $\frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2y = 0$ has as solutions $y_1(x) = e^{2x}$ and $y_2(x) = e^x$. Any linear combination $y(x) = c_1y_1(x) + c_2y_2(x)$ is a solution of the differential equation, and so $\text{Span}(e^{2x}, e^x)$ is contained in the set of solutions of the differential equation (called the solution space for the differential equation). In fact, the solution space is spanned by e^{2x} and e^x, and so is a subspace of the vector space of functions. In general, the solution space for a homogeneous differential equation is a vector space, meaning that any linear combination of solutions is again a solution.

2.2 Basis and Dimension of a Vector Space

Let V be a vector space over a field F.

Definitions:

A set of vectors $\mathcal{B}$ in a vector space V is a **basis** for V if

- $\mathcal{B}$ is a linearly independent set, and
- $\text{Span}(\mathcal{B}) = V$.

The set $\mathcal{E}_n = \left\{ \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \right\}$ is the **standard basis** for F^n.

The number of vectors in a basis for a vector space V is the **dimension** of V, denoted by $\dim(V)$. If a basis for V contains a finite number of vectors, then V is **finite dimensional**. Otherwise, V is **infinite dimensional**, and we write $\dim(V) = \infty$.

Facts: All the following facts, except those with a specific reference, can be found in [Lay03, Sections 4.3 and 4.5].

1. Every vector space has a basis.
2. The standard basis for F^n is a basis for F^n, and so $\dim F^n = n$.
3. A basis $\mathcal{B}$ in a vector space V is the largest set of linearly independent vectors in V that contains $\mathcal{B}$, and it is the smallest set of vectors in V that contains $\mathcal{B}$ and spans V.
4. The empty set is a basis for the trivial vector space $\{\mathbf{0}\}$, and $\dim(\{\mathbf{0}\}) = 0$.
5. If the set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ spans a vector space V, then some subset of S forms a basis for V. In particular, if one of the vectors, say $\mathbf{v}_i$, is a linear combination of the remaining vectors, then the set formed from S by removing $\mathbf{v}_i$ will be "closer" to a basis for V. This process can be continued until the remaining vectors form a basis for V.
6. If S is a linearly independent set in a vector space V, then S can be expanded, if necessary, to a basis for V.
7. No nontrivial vector space over a field with more than two elements has a unique basis.
8. If a vector space V has a basis containing n vectors, then every basis of V must contain n vectors. Similarly, if V has an infinite basis, then every basis of V must be infinite. So the dimension of V is unique.
9. Let $\dim(V) = n$ and let S be a set containing n vectors. The following are equivalent:
    - S is a basis for V.
    - S spans V.
    - S is linearly independent.

10. If $\dim(V) = n$, then any subset of V containing more than n vectors is linearly dependent.
11. If $\dim(V) = n$, then any subset of V containing fewer than n vectors does not span V.
12. [Lay03, Section 4.4] If $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ is a basis for a vector space V, then each $\mathbf{x} \in V$ can be expressed as a unique linear combination of the vectors in $\mathcal{B}$. That is, for each $\mathbf{x} \in V$ there is a unique set of scalars $c_1, c_2, \ldots, c_p$ such that $\mathbf{x} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \cdots + c_p\mathbf{b}_p$.

Examples:

1. In $\mathbb{R}^2$, $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$ are linearly independent, and they span $\mathbb{R}^2$. So they form a basis for $\mathbb{R}^2$ and $\dim(\mathbb{R}^2) = 2$.
2. In $F[x]$, the set $\{1, x, x^2, \ldots, x^n\}$ is a basis for $F[x; n]$ for any $n \in \mathbb{N}$. The infinite set $\{1, x, x^2, x^3, \ldots\}$ is a basis for $F[x]$, meaning $\dim(F[x]) = \infty$.
3. The set of $m \times n$ matrices E_{ij} having a 1 in the i, j-entry and zeros everywhere else forms a basis for $F^{m\times n}$. Since there are mn such matrices, $\dim(F^{m\times n}) = mn$.
4. The set $S = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ clearly spans $\mathbb{R}^2$, but it is not a linearly independent set. However, removing any single vector from S will cause the remaining vectors to be a basis for $\mathbb{R}^2$, because any pair of vectors is linearly independent and still spans $\mathbb{R}^2$.
5. The set $S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ is linearly independent, but it cannot be a basis for $\mathbb{R}^4$ since it does not span $\mathbb{R}^4$. However, we can start expanding it to a basis for $\mathbb{R}^4$ by first adding a vector that is not in the span of S, such as $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$. Then since these three vectors still do not span $\mathbb{R}^4$, we can add a vector that is not in their span, such as $\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$. These four vectors now span $\mathbb{R}^4$ and they are linearly independent, so they form a basis for $\mathbb{R}^4$.
6. Additional techniques for determining whether a given finite set of vectors is linearly independent or spans a given subspace can be found in Sections 2.5 and 2.6.

Applications:

1. Because $y_1(x) = e^{2x}$ and $y_2(x) = e^x$ are linearly independent and span the solution space for the homogeneous differential equation $\frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2y = 0$, they form a basis for the solution space and the solution space has dimension 2.

2.3 Direct Sum Decompositions

Throughout this section, V will be a vector space over a field F, and W_i, for $i = 1, \ldots, k$, will be subspaces of V. For facts and general reading for this section, see [HK71].

Definitions:

The **sum** of subspaces W_i, for $i = 1, \ldots, k$, is $\sum_{i=1}^{k} W_i = W_1 + \cdots + W_k = \{\mathbf{w}_1 + \cdots + \mathbf{w}_k \mid \mathbf{w}_i \in W_i\}$. The sum $W_1 + \cdots + W_k$ is a **direct sum** if for all $i = 1, \ldots, k$, we have $W_i \cap \sum_{j \neq i} W_j = \{\mathbf{0}\}$. $W = W_1 \oplus \cdots \oplus W_k$ denotes that $W = W_1 + \cdots + W_k$ and the sum is direct. The subspaces W_i, for $i = i, \ldots, k$, are **independent** if for $\mathbf{w}_i \in W_i$, $\mathbf{w}_1 + \cdots + \mathbf{w}_k = \mathbf{0}$ implies $\mathbf{w}_i = \mathbf{0}$ for all $i = 1, \ldots, k$. Let V_i, for $i = 1, \ldots, k$, be vector spaces over F. The **external direct sum** of the V_i, denoted $V_1 \times \cdots \times V_k$, is the cartesian product of V_i, for $i = 1, \ldots, k$, with coordinate-wise operations. Let W be a subspace of V. An **additive coset** of W is a subset of the form $v + W = \{v + w \mid w \in W\}$ with $v \in V$. The **quotient** of V by W, denoted V/W, is the set of additive cosets of W with operations $(v_1 + W) + (v_2 + W) = (v_1 + v_2) + W$ and $c(v + W) = (cv) + W$, for any $c \in F$. Let $V = W \oplus U$, let $\mathcal{B}_W$ and $\mathcal{B}_U$ be bases for W and U respectively, and let $\mathcal{B} = \mathcal{B}_W \cup \mathcal{B}_U$. The **induced basis** of $\mathcal{B}$ in V/W is the set of vectors $\{u + W \mid u \in \mathcal{B}_U\}$.

Facts:

1. $W = W_1 \oplus W_2$ if and only if $W = W_1 + W_2$ and $W_1 \cap W_2 = \{\mathbf{0}\}$.
2. If W is a subspace of V, then there exists a subspace U of V such that $V = W \oplus U$. Note that U is not usually unique.
3. Let $W = W_1 + \cdots + W_k$. The following are equivalent:
   - $W = W_1 \oplus \cdots \oplus W_k$. That is, for all $i = 1, \ldots, k$, we have $W_i \cap \sum_{j \neq i} W_j = \{\mathbf{0}\}$.
   - $W_i \cap \sum_{j=1}^{i-1} W_j = \{\mathbf{0}\}$, for all $i = 2, \ldots, k$.
   - For each $\mathbf{w} \in W$, $\mathbf{w}$ can be expressed in exactly one way as a sum of vectors in $W_1, \ldots, W_k$. That is, there exist unique $\mathbf{w}_i \in W_i$, such that $\mathbf{w} = \mathbf{w}_1 + \cdots + \mathbf{w}_k$.
   - The subspaces W_i, for $i = 1, \ldots, k$, are independent.
   - If $\mathcal{B}_i$ is an (ordered) basis for W_i, then $\mathcal{B} = \bigcup_{i=1}^{k} \mathcal{B}_i$ is an (ordered) basis for W.
4. If $\mathcal{B}$ is a basis for V and $\mathcal{B}$ is partitioned into disjoint subsets $\mathcal{B}_i$, for $i = 1, \ldots, k$, then $V = \mathrm{Span}(\mathcal{B}_1) \oplus \cdots \oplus \mathrm{Span}(\mathcal{B}_k)$.
5. If S is a linearly independent subset of V and S is partitioned into disjoint subsets S_i, for $i = 1, \ldots, k$, then the subspaces $\mathrm{Span}(S_1), \ldots, \mathrm{Span}(S_k)$ are independent.
6. If V is finite dimensional and $V = W_1 + \cdots + W_k$, then $\dim(V) = \dim(W_1) + \cdots + \dim(W_k)$ if and only if $V = W_1 \oplus \cdots \oplus W_k$.
7. Let V_i, for $i = 1, \ldots, k$, be vector spaces over F.
   - $V_1 \times \cdots \times V_k$ is a vector space over F.
   - $\widehat{V}_i = \{(0, \ldots, 0, v_i, 0, \ldots, 0) \mid v_i \in V_i\}$ (where v_i is the ith coordinate) is a subspace of $V_1 \times \cdots \times V_k$.
   - $V_1 \times \cdots \times V_k = \widehat{V}_1 \oplus \cdots \oplus \widehat{V}_k$.
   - If V_i, for $i = 1, \ldots, k$, are finite dimensional, then $\dim \widehat{V}_i = \dim V_i$ and $\dim(V_1 \times \cdots \times V_k) = \dim V_1 + \cdots + \dim V_k$.
8. If W is a subspace of V, then the quotient V/W is a vector space over F.
9. Let $V = W \oplus U$, let $\mathcal{B}_W$ and $\mathcal{B}_U$ be bases for W and U respectively, and let $\mathcal{B} = \mathcal{B}_W \cup \mathcal{B}_U$. The induced basis of $\mathcal{B}$ in V/W is a basis for V/W and $\dim(V/W) = \dim U$.

Examples:

1. Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for V. Then $V = \mathrm{Span}(\mathbf{v}_1) \oplus \cdots \oplus \mathrm{Span}(\mathbf{v}_n)$.
2. Let $X = \left\{ \begin{bmatrix} x \\ 0 \end{bmatrix} \mid x \in \mathbb{R} \right\}$, $Y = \left\{ \begin{bmatrix} 0 \\ y \end{bmatrix} \mid y \in \mathbb{R} \right\}$, and $Z = \left\{ \begin{bmatrix} z \\ z \end{bmatrix} \mid z \in \mathbb{R} \right\}$. Then $\mathbb{R}^2 = X \oplus Y = Y \oplus Z = X \oplus Z$.

3. In $F^{n\times n}$, let W_1 be the subspace of symmetric matrices and W_2 be the subspace of skew-symmetric matrices. Clearly, $W_1 \cap W_2 = \{\mathbf{0}\}$. For any $A \in F^{n\times n}$, $A = \frac{A + A^T}{2} + \frac{A - A^T}{2}$, where $\frac{A + A^T}{2} \in W_1$ and $\frac{A - A^T}{2} \in W_2$. Therefore, $F^{n\times n} = W_1 \oplus W_2$.
4. Recall that the function $f \in \mathcal{C}(\mathbb{R})$ is even if $f(-x) = f(x)$ for all x, and f is odd if $f(-x) = -f(x)$ for all x. Let W_1 be the subspace of even functions and W_2 be the subspace of odd functions. Clearly, $W_1 \cap W_2 = \{\mathbf{0}\}$. For any $f \in \mathcal{C}(\mathbb{R})$, $f = f_1 + f_2$, where $f_1(x) = \frac{f(x) + f(-x)}{2} \in W_1$ and $f_1(x) = \frac{f(x) - f(-x)}{2} \in W_2$. Therefore, $\mathcal{C}(\mathbb{R}) = W_1 \oplus W_2$.
5. Given a subspace W of V, we can find a subspace U such that $V = W \oplus U$ by choosing a basis for W, extending this linearly independent set to a basis for V, and setting U equal to the span of the basis vectors not in W. For example, in $\mathbb{R}^3$, Let $W = \left\{ \begin{bmatrix} a \\ -2a \\ a \end{bmatrix} \mid a \in \mathbb{R} \right\}$. If $\mathbf{w} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$, then $\{\mathbf{w}\}$ is a basis for W. Extend this to a basis for $\mathbb{R}^3$, for example by adjoining $\mathbf{e}_1$ and $\mathbf{e}_2$. Thus, $V = W \oplus U$, where $U = \text{Span}(\mathbf{e}_1, \mathbf{e}_2)$. Note: there are many other ways to extend the basis, and many other possible U.
6. In the external direct sum $\mathbb{R}[x; 2] \times \mathbb{R}^{2\times 2}$, $\left(2x^2 + 7, \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\right) + 3\left(x^2 + 4x - 2, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right) = \left(5x^2 + 12x + 1, \begin{bmatrix} 1 & 5 \\ 0 & 4 \end{bmatrix}\right)$.
7. The subspaces X, Y, Z of $\mathbb{R}^2$ in Example 2 have bases $\mathcal{B}_X = \left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$, $\mathcal{B}_Y = \left\{\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$, $\mathcal{B}_Z = \left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$, respectively. Then $\mathcal{B}_{XY} = \mathcal{B}_X \cup \mathcal{B}_Y$ and $\mathcal{B}_{XZ} = \mathcal{B}_X \cup \mathcal{B}_Z$ are bases for $\mathbb{R}^2$. In $\mathbb{R}^2/X$, the induced bases of $\mathcal{B}_{XY}$ and $\mathcal{B}_{XZ}$ are $\left\{\begin{bmatrix} 0 \\ 1 \end{bmatrix} + X\right\}$ and $\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix} + X\right\}$, respectively. These are equal because $\begin{bmatrix} 1 \\ 1 \end{bmatrix} + X = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} + X = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + X$.

2.4 Matrix Range, Null Space, Rank, and the Dimension Theorem

Definitions:

For any matrix $A \in F^{m\times n}$, the **range** of A, denoted by range(A), is the set of all linear combinations of the columns of A. If $A = [\mathbf{m}_1\ \mathbf{m}_2\ \dots\ \mathbf{m}_n]$, then range($A$) $=$ Span($\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_n$). The range of A is also called the **column space** of A.

The **row space** of A, denoted by RS(A), is the set of all linear combinations of the rows of A. If $A = [\mathbf{v}_1\ \mathbf{v}_2\ \dots\ \mathbf{v}_m]^T$, then RS($A$) $=$ Span($\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$).

The **kernel** of A, denoted by ker(A), is the set of all solutions to the homogeneous equation $A\mathbf{x} = \mathbf{0}$. The kernel of A is also called the **null space** of A, and its dimension is called the **nullity** of A, denoted by null(A).

The **rank** of A, denoted by rank(A), is the number of leading entries in the reduced row echelon form of A (or any row echelon form of A). (See Section 1.3 for more information.)

$A, B \in F^{m\times n}$ are **equivalent** if $B = C_1^{-1}AC_2$ for some invertible matrices $C_1 \in F^{m\times m}$ and $C_2 \in F^{n\times n}$. $A, B \in F^{n\times n}$ are **similar** if $B = C^{-1}AC$ for some invertible matrix $C \in F^{n\times n}$. For square matrices $A_1 \in F^{n_1\times n_1}, \ldots, A_k \in F^{n_k\times n_k}$, the **matrix direct sum** $A = A_1 \oplus \cdots \oplus A_k$ is the block diagonal matrix with the matrices A_i down the diagonal. That is, $A = \begin{bmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_k \end{bmatrix}$, where $A \in F^{n\times n}$ with $n = \sum_{i=1}^{k} n_i$.

Facts: Unless specified otherwise, the following facts can be found in [Lay03, Sections 2.8, 4.2, 4.5, and 4.6].

1. The range of an $m \times n$ matrix A is a subspace of F^m.
2. The columns of A corresponding to the pivot columns in the reduced row echelon form of A (or any row echelon form of A) give a basis for range(A). Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in F^m$. If matrix $A = [\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_k]$, then a basis for range$(A)$ will be a linearly independent subset of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ having the same span.
3. dim(range(A)) = rank(A).
4. The kernel of an $m \times n$ matrix A is a subspace of F^n.
5. If the reduced row echelon form of A (or any row echelon form of A) has k pivot columns, then null$(A) = n - k$.
6. If two matrices A and B are row equivalent, then RS(A) = RS(B).
7. The row space of an $m \times n$ matrix A is a subspace of F^n.
8. The pivot rows in the reduced row echelon form of A (or any row echelon form of A) give a basis for RS(A).
9. dim(RS(A)) = rank(A).
10. rank(A) = rank(A^T).
11. (Dimension Theorem) For any $A \in F^{m\times n}$, $n = \text{rank}(A) + \text{null}(A)$. Similarly, $m = \dim(\text{RS}(A)) + \text{null}(A^T)$.
12. A vector $\mathbf{b} \in F^m$ is in range(A) if and only if the equation $A\mathbf{x} = \mathbf{b}$ has a solution. So range$(A) = F^m$ if and only if the equation $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b} \in F^m$.
13. A vector $\mathbf{a} \in F^n$ is in RS(A) if and only if the equation $A^T\mathbf{y} = \mathbf{a}$ has a solution. So RS$(A) = F^n$ if and only if the equation $A^T\mathbf{y} = \mathbf{a}$ has a solution for every $\mathbf{a} \in F^n$.
14. If $\mathbf{a}$ is a solution to the equation $A\mathbf{x} = \mathbf{b}$, then $\mathbf{a} + \mathbf{v}$ is also a solution for any $\mathbf{v} \in \ker(A)$.
15. [HJ85, p. 14] If $A \in F^{m\times n}$ is rank 1, then there are vectors $\mathbf{v} \in F^m$ and $\mathbf{u} \in F^n$ so that $A = \mathbf{v}\mathbf{u}^T$.
16. If $A \in F^{m\times n}$ is rank k, then A is a sum of k rank 1 matrices. That is, there exist $A_1, \ldots, A_k$ with $A = A_1 + \cdots + A_k$ and rank$(A_i) = 1$, for $i = 1, \ldots, k$.
17. [HJ85, p. 13] The following are all equivalent statements about a matrix $A \in F^{m\times n}$.
    (a) The rank of A is k.
    (b) dim(range(A)) $= k$.
    (c) The reduced row echelon form of A has k pivot columns.
    (d) A row echelon form of A has k pivot columns.
    (e) The largest number of linearly independent columns of A is k.
    (f) The largest number of linearly independent rows of A is k.
18. [HJ85, p. 13] (Rank Inequalities) (Unless specified otherwise, assume that $A, B \in F^{m\times n}$.)
    (a) rank$(A) \leq \min(m, n)$.
    (b) If a new matrix B is created by deleting rows and/or columns of matrix A, then rank$(B) \leq$ rank(A).
    (c) rank$(A + B) \leq$ rank(A) + rank(B).
    (d) If A has a $p \times q$ submatrix of 0s, then rank$(A) \leq (m - p) + (n - q)$.

(e) If $A \in F^{m\times k}$ and $B \in F^{k\times n}$, then

$$\operatorname{rank}(A) + \operatorname{rank}(B) - k \le \operatorname{rank}(AB) \le \min\{\operatorname{rank}(A), \operatorname{rank}(B)\}.$$

19. [HJ85, pp. 13–14] (Rank Equalities)
    (a) If $A \in \mathbb{C}^{m\times n}$, then $\operatorname{rank}(A^*) = \operatorname{rank}(A^T) = \operatorname{rank}(\overline{A}) = \operatorname{rank}(A)$.
    (b) If $A \in \mathbb{C}^{m\times n}$, then $\operatorname{rank}(A^*A) = \operatorname{rank}(A)$. If $A \in \mathbb{R}^{m\times n}$, then $\operatorname{rank}(A^TA) = \operatorname{rank}(A)$.
    (c) Rank is unchanged by left or right multiplication by a nonsingular matrix. That is, if $A \in F^{n\times n}$ and $B \in F^{m\times m}$ are nonsingular, and $M \in F^{m\times n}$, then

    $$\operatorname{rank}(AM) = \operatorname{rank}(M) = \operatorname{rank}(MB) = \operatorname{rank}(AMB).$$

    (d) If $A, B \in F^{m\times n}$, then $\operatorname{rank}(A) = \operatorname{rank}(B)$ if and only if there exist nonsingular matrices $X \in F^{m\times m}$ and $Y \in F^{n\times n}$ such that $A = XBY$ (i.e., if and only if A is equivalent to B).
    (e) If $A \in F^{m\times n}$ has rank k, then $A = XBY$, for some $X \in F^{m\times k}$, $Y \in F^{k\times n}$, and nonsingular $B \in F^{k\times k}$.
    (f) If $A_1 \in F^{n_1\times n_1}, \ldots, A_k \in F^{n_k\times n_k}$, then $\operatorname{rank}(A_1 \oplus \cdots \oplus A_k) = \operatorname{rank}(A_1) + \cdots + \operatorname{rank}(A_k)$.
20. Let $A, B \in F^{n\times n}$ with A similar to B.
    (a) A is equivalent to B.
    (b) $\operatorname{rank}(A) = \operatorname{rank}(B)$.
    (c) $\operatorname{tr} A = \operatorname{tr} B$.
21. Equivalence of matrices is an equivalence relation on $F^{m\times n}$.
22. Similarity of matrices is an equivalence relation on $F^{n\times n}$.
23. If $A \in F^{m\times n}$ and $\operatorname{rank}(A) = k$, then A is equivalent to $\begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}$, and so any two matrices of the same size and rank are equivalent.
24. (For information on the determination of whether two matrices are similar, see Chapter 6.)
25. [Lay03, Sec. 6.1] If $A \in \mathbb{R}^{n\times n}$, then for any $\mathbf{x} \in \operatorname{RS}(A)$ and any $\mathbf{y} \in \ker(A)$, $\mathbf{x}^T\mathbf{y} = 0$. So the row space and kernel of a real matrix are orthogonal to one another. (See Chapter 5 for more on orthogonality.)

Examples:

1. If $A = \begin{bmatrix} 1 & 7 & -2 \\ 0 & -1 & 1 \\ 2 & 13 & -3 \end{bmatrix} \in \mathbb{R}^{3\times 3}$, then any vector of the form $\begin{bmatrix} a + 7b - 2c \\ -b + c \\ 2a + 13b - 3c \end{bmatrix} \left(= \begin{bmatrix} 1 & 7 & -2 \\ 0 & -1 & 1 \\ 2 & 13 & -3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \right)$ is in $\operatorname{range}(A)$, for any $a, b, c \in \mathbb{R}$. Since a row echelon form of A is $\begin{bmatrix} 1 & 7 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$, we know that the set $\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 7 \\ -1 \\ 13 \end{bmatrix} \right\}$ is a basis for $\operatorname{range}(A)$, and the set $\left\{ \begin{bmatrix} 1 \\ 7 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}$ is a basis for $\operatorname{RS}(A)$. Since its reduced row echelon form is $\begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$, the set $\left\{ \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}$ is another basis for $\operatorname{RS}(A)$.

2. If $A = \begin{bmatrix} 1 & 7 & -2 \\ 0 & -1 & 1 \\ 2 & 13 & -3 \end{bmatrix} \in \mathbb{R}^{3\times 3}$, then using the reduced row echelon form given in the previous example, solutions to $A\mathbf{x} = \mathbf{0}$ have the form $\mathbf{x} = c\begin{bmatrix} -5 \\ 1 \\ 1 \end{bmatrix}$, for any $c \in \mathbb{R}$. So $\ker(A) = \operatorname{Span}\left(\begin{bmatrix} -5 \\ 1 \\ 1 \end{bmatrix}\right)$.

3. If $A \in \mathbb{R}^{3\times 5}$ has the reduced row echelon form $\begin{bmatrix} 1 & 0 & 3 & 0 & 2 \\ 0 & 1 & -2 & 0 & 7 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}$, then any solution to $A\mathbf{x} = \mathbf{0}$ has the form

$$\mathbf{x} = c_1\begin{bmatrix} -3 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + c_2\begin{bmatrix} -2 \\ -7 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

for some $c_1, c_2 \in \mathbb{R}$. So,

$$\ker(A) = \operatorname{Span}\left(\begin{bmatrix} -3 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ -7 \\ 0 \\ 1 \\ 1 \end{bmatrix}\right).$$

4. Example 1 above shows that $\left\{\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 7 \\ -1 \\ 13 \end{bmatrix}\right\}$ is a linearly independent set having the same span as the set $\left\{\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 7 \\ -1 \\ 13 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix}\right\}$.

5. $\begin{bmatrix} 1 & 7 \\ 2 & -3 \end{bmatrix}$ is similar to $\begin{bmatrix} 37 & -46 \\ 31 & -39 \end{bmatrix}$ because $\begin{bmatrix} 37 & -46 \\ 31 & -39 \end{bmatrix} = \begin{bmatrix} -2 & 3 \\ 3 & -4 \end{bmatrix}^{-1}\begin{bmatrix} 1 & 7 \\ 2 & -3 \end{bmatrix}\begin{bmatrix} -2 & 3 \\ 3 & -4 \end{bmatrix}$.

2.5 Nonsingularity Characterizations

From the previous discussion, we can add to the list of nonsingularity characterizations of a square matrix that was started in the previous chapter.

Facts: The following facts can be found in [HJ85, p. 14] or [Lay03, Sections 2.3 and 4.6].

1. If $A \in F^{n\times n}$, then the following are equivalent.
   (a) A is nonsingular.
   (b) The columns of A are linearly independent.
   (c) The dimension of range(A) is n.

(d) The range of A is F^n.
(e) The equation $A\mathbf{x} = \mathbf{b}$ is consistent for each $\mathbf{b} \in F^n$.
(f) If the equation $A\mathbf{x} = \mathbf{b}$ is consistent, then the solution is unique.
(g) The equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for each $\mathbf{b} \in F^n$.
(h) The rows of A are linearly independent.
(i) The dimension of $\mathrm{RS}(A)$ is n.
(j) The row space of A is F^n.
(k) The dimension of $\ker(A)$ is 0.
(l) The only solution to $A\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$.
(m) The rank of A is n.
(n) The determinant of A is nonzero. (See Section 4.1 for the definition of the determinant.)

2.6 Coordinates and Change of Basis

Coordinates are used to transform a problem in a more abstract vector space (e.g., the vector space of polynomials of degree less than or equal to 3) to a problem in F^n.

Definitions:

Suppose that $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ is an ordered basis for a vector space V over a field F and $\mathbf{x} \in V$. The **coordinates of x relative to the ordered basis** $\mathcal{B}$ (or the $\mathcal{B}$**-coordinates of x**) are the scalar coefficients $c_1, c_2, \ldots, c_n \in F$ such that $\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n$. Whenever coordinates are involved, the vector space is assumed to be nonzero and finite dimensional.

If $c_1, c_2, \ldots, c_n$ are the $\mathcal{B}$-coordinates of $\mathbf{x}$, then the vector in F^n,

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix},$$

is the **coordinate vector of x relative to** $\mathcal{B}$ or the $\mathcal{B}$**-coordinate vector of x**.

The mapping $\mathbf{x} \to [\mathbf{x}]_{\mathcal{B}}$ is the **coordinate mapping determined by** $\mathcal{B}$.

If $\mathcal{B}$ and $\mathcal{B}'$ are ordered bases for the vector space F^n, then the **change-of-basis matrix** from $\mathcal{B}$ to $\mathcal{B}'$ is the matrix whose columns are the $\mathcal{B}'$-coordinate vectors of the vectors in $\mathcal{B}$ and is denoted by ${}_{\mathcal{B}'}[I]_{\mathcal{B}}$. Such a matrix is also called a **transition matrix**.

Facts: The following facts can be found in [Lay03, Sections 4.4 and 4.7] or [HJ85, Section 0.10]:

1. For any vector $\mathbf{x} \in F^n$ with the standard ordered basis $\mathcal{E}_n = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$, we have $\mathbf{x} = [\mathbf{x}]_{\mathcal{E}_n}$.
2. For any ordered basis $\mathcal{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ of a vector space V, we have $[\mathbf{b}_i]_{\mathcal{B}} = \mathbf{e}_i$.
3. If $\dim(V) = n$, then the coordinate mapping is a one-to-one linear transformation from V onto F^n. (See Chapter 3 for the definition of linear transformation.)
4. If $\mathcal{B}$ is an ordered basis for a vector space V and $\mathbf{v}_1, \mathbf{v}_2 \in V$, then $\mathbf{v}_1 = \mathbf{v}_2$ if and only if $[\mathbf{v}_1]_{\mathcal{B}} = [\mathbf{v}_2]_{\mathcal{B}}$.
5. Let V be a vector space over a field F, and suppose $\mathcal{B}$ is an ordered basis for V. Then for any $\mathbf{x}, \mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ and $c_1, \ldots, c_k \in F$, $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ if and only if $[\mathbf{x}]_{\mathcal{B}} = c_1[\mathbf{v}_1]_{\mathcal{B}} + \cdots + c_k[\mathbf{v}_k]_{\mathcal{B}}$. So, for any $\mathbf{x}, \mathbf{v}_1, \ldots, \mathbf{v}_k \in V$, $\mathbf{x} \in \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ if and only if $[\mathbf{x}]_{\mathcal{B}} \in \mathrm{Span}([\mathbf{v}_1]_{\mathcal{B}}, \ldots, [\mathbf{v}_k]_{\mathcal{B}})$.
6. Suppose $\mathcal{B}$ is an ordered basis for an n-dimensional vector space V over a field F and $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$. The set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly independent in V if and only if the set $S' = \{[\mathbf{v}_1]_{\mathcal{B}}, \ldots, [\mathbf{v}_k]_{\mathcal{B}}\}$ is linearly independent in F^n.

7. Let V be a vector space over a field F with $\dim(V) = n$, and suppose $\mathcal{B}$ is an ordered basis for V. Then $\text{Span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k) = V$ for some $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in V$ if and only if $\text{Span}([\mathbf{v}_1]_{\mathcal{B}}, [\mathbf{v}_2]_{\mathcal{B}}, \ldots, [\mathbf{v}_k]_{\mathcal{B}}) = F^n$.
8. Suppose $\mathcal{B}$ is an ordered basis for a vector space V over a field F with $\dim(V) = n$, and let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a subset of V. Then S is a basis for V if and only if $\{[\mathbf{v}_1]_{\mathcal{B}}, \ldots, [\mathbf{v}_n]_{\mathcal{B}}\}$ is a basis for F^n if and only if the matrix $[[\mathbf{v}_1]_{\mathcal{B}}, \ldots, [\mathbf{v}_n]_{\mathcal{B}}]$ is invertible.
9. If $\mathcal{B}$ and $\mathcal{B}'$ are ordered bases for a vector space V, then $[\mathbf{x}]_{\mathcal{B}'} = {}_{\mathcal{B}'}[I]_{\mathcal{B}}\, [\mathbf{x}]_{\mathcal{B}}$ for any $\mathbf{x} \in V$. Furthermore, ${}_{\mathcal{B}'}[I]_{\mathcal{B}}$ is the only matrix such that for any $\mathbf{x} \in V$, $[\mathbf{x}]_{\mathcal{B}'} = {}_{\mathcal{B}'}[I]_{\mathcal{B}}\, [\mathbf{x}]_{\mathcal{B}}$.
10. Any change-of-basis matrix is invertible.
11. If B is invertible, then B is a change-of-basis matrix. Specifically, if $B = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_n] \in F^{n\times n}$, then $B = {}_{\mathcal{E}_n}[I]_{\mathcal{B}}$, where $\mathcal{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ is an ordered basis for F^n.
12. If $\mathcal{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ is an ordered basis for F^n, then ${}_{\mathcal{E}_n}[I]_{\mathcal{B}} = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_n]$.
13. If $\mathcal{B}$ and $\mathcal{B}'$ are ordered bases for a vector space V, then ${}_{\mathcal{B}}[I]_{\mathcal{B}'} = ({}_{\mathcal{B}'}[I]_{\mathcal{B}})^{-1}$.
14. If $\mathcal{B}$ and $\mathcal{B}'$ are ordered bases for F^n, then ${}_{\mathcal{B}'}[I]_{\mathcal{B}} = ({}_{\mathcal{B}'}[I]_{\mathcal{E}_n})({}_{\mathcal{E}_n}[I]_{\mathcal{B}})$.

Examples:

1. If $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in F[x; n]$ with the standard ordered basis $\mathcal{B} = (1, x, x^2, \ldots, x^n)$, then $[p(x)]_{\mathcal{B}} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}$.
2. The set $\mathcal{B} = \left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix} \right)$ forms an ordered basis for $\mathbb{R}^2$. If $\mathcal{E}_2$ is the standard ordered basis for $\mathbb{R}^2$, then the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{E}_2$ is ${}_{\mathcal{E}_2}[T]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ -1 & 3 \end{bmatrix}$, and $({}_{\mathcal{E}_2}[T]_{\mathcal{B}})^{-1} = \begin{bmatrix} 1 & 0 \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}$. So for $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ in the standard ordered basis, we find that $[\mathbf{v}]_{\mathcal{B}} = ({}_{\mathcal{E}_2}[T]_{\mathcal{B}})^{-1}\mathbf{v} = \begin{bmatrix} 3 \\ \frac{4}{3} \end{bmatrix}$. To check this, we can easily see that $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \frac{4}{3} \begin{bmatrix} 0 \\ 3 \end{bmatrix}$.
3. The set $\mathcal{B}' = (1, 1+x, 1+x^2)$ is an ordered basis for $\mathbb{R}[x; 2]$, and using the standard ordered basis $\mathcal{B} = (1, x, x^2)$ for $\mathbb{R}[x; 2]$ we have ${}_{\mathcal{B}}[P]_{\mathcal{B}'} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. So, $({}_{\mathcal{B}}[P]_{\mathcal{B}'})^{-1} = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $[5 - 2x + 3x^2]_{\mathcal{B}'} = ({}_{\mathcal{B}}[P]_{\mathcal{B}'})^{-1} \begin{bmatrix} 5 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix}$. Of course, we can see $5 - 2x + 3x^2 = 4(1) - 2(1+x) + 3(1+x^2)$.
4. If we want to change from the ordered basis $\mathcal{B}_1 = \left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix} \right)$ in $\mathbb{R}^2$ to the ordered basis $\mathcal{B}_2 = \left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \end{bmatrix} \right)$, then the resulting change-of-basis matrix is ${}_{\mathcal{B}_2}[T]_{\mathcal{B}_1} = ({}_{\mathcal{E}_2}[T]_{\mathcal{B}_2})^{-1}({}_{\mathcal{E}_2}[T]_{\mathcal{B}_1}) = \begin{bmatrix} 2 & 5 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} -1 & 3 \\ \frac{3}{5} & -\frac{6}{5} \end{bmatrix}$.

5. Let $S = \{5 - 2x + 3x^2, 3 - x + 2x^2, 8 + 3x\}$ in $\mathbb{R}[x; 2]$ with the standard ordered basis $\mathcal{B} = (1, x, x^2)$. The matrix $A = \begin{bmatrix} 5 & 3 & 8 \\ -2 & -1 & 3 \\ 3 & 2 & 0 \end{bmatrix}$ contains the $\mathcal{B}$-coordinate vectors for the polynomials in S and it has row echelon form $\begin{bmatrix} 5 & 3 & 8 \\ 0 & 1 & 31 \\ 0 & 0 & 1 \end{bmatrix}$. Since this row echelon form shows that A is nonsingular, we know by Fact 8 above that S is a basis for $\mathbb{R}[x; 2]$.

2.7 Idempotence and Nilpotence

Definitions:

A is an **idempotent** if $A^2 = A$.

A is **nilpotent** if, for some $k \geq 0$, $A^k = 0$.

Facts: All of the following facts except those with a specific reference are immediate from the definitions.

1. Every idempotent except the identity matrix is singular.
2. Let $A \in F^{n \times n}$. The following statements are equivalent.
   (a) A is an idempotent.
   (b) $I - A$ is an idempotent.
   (c) If $\mathbf{v} \in \text{range}(A)$, then $A\mathbf{v} = \mathbf{v}$.
   (d) $F^n = \ker A \oplus \text{range} A$.
   (e) [HJ85, p. 37 and p. 148] A is similar to $\begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}$, for some $k \leq n$.
3. If A_1 and A_2 are idempotents of the same size and commute, then $A_1 A_2$ is an idempotent.
4. If A_1 and A_2 are idempotents of the same size and $A_1 A_2 = A_2 A_1 = 0$, then $A_1 + A_2$ is an idempotent.
5. If $A \in F^{n \times n}$ is nilpotent, then $A^n = 0$.
6. If A is nilpotent and B is of the same size and commutes with A, then AB is nilpotent.
7. If A_1 and A_2 are nilpotent matrices of the same size and $A_1 A_2 = A_2 A_1 = 0$, then $A_1 + A_2$ is nilpotent.

Examples:

1. $\begin{bmatrix} -8 & 12 \\ -6 & 9 \end{bmatrix}$ is an idempotent. $\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$ is nilpotent.

References

[Lay03] D. C. Lay. *Linear Algebra and Its Applications*, 3rd ed. Addison-Wesley, Reading, MA, 2003.

[HK71] K. H. Hoffman and R. Kunze. *Linear Algebra*, 2nd ed. Prentice-Hall, Upper Saddle River, NJ, 1971.

[HJ85] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.