Espacios vectoriales

Credit: This notes are 100% from chapter 6 of the book entitled *Linear Algebra. A Modern Introduction* by David Poole. Thomson. Australia. 2006.

Vector Spaces and Subspaces

In this section we define generalized "vectors" that arise in a wide variety of examples.

Vector space: Let V be a set on which two operations, called addition and scalar multiplication (by scalar we will usually mean the Real Numbers), have been defined. If $\bar{u}$ and $\bar{v}$ are in V, the sum of $\bar{u}$ and $\bar{v}$ is denoted by $\bar{u}+\bar{v}$, and if c is a scalar, the scalar multiple of $\bar{u}$ by c is denoted by $c\bar{u}$. If the following axioms hold for all $\bar{u}$, $\bar{v}$ and $\bar{w}$ in V and for all scalars c and d, then V is called a vector space and its elements are called vectors.

1. $\bar{u}+\bar{v}$ is in V. Closure under addition
2. $\bar{u}+\bar{v}$=$\bar{v}+\bar{u}$. Commutativity
3. $(\bar{u}+\bar{v})+\bar{w}=\bar{u}+(\bar{v}+\bar{w})$. Associativity
4. There exists and element $\bar{0}$ in V, called a zero vector, such that $\bar{u}+\bar{0}=\bar{u}$.
5. For each $\bar{u}$ in V, there is an element - $\bar{u}$ in V such that $\bar{u}+(-\bar{u})=\bar{0}$.
6. $c\bar{u}$ is in V. Closure under scalar multiplication
7. $c(\bar{u}+\bar{v})=c\bar{u}+c\bar{v}$. Distributivity
8. $(c+d)\bar{u}=c\bar{u}+d\bar{u}$. Distributivity
9. $c(d\bar{u}=(cd)\bar{u}$
10. $1\bar{u}=\bar{u}$

Comments:

- By "scalar" we will usually mean the real numbers. In this case we refer to V as a real vector space (or a vector space over the real numbers). It is also possible for scalars to be complex numbers or to belong to $\mathbb{Z}_p$, where p is prime. In these cases, V is called a complex vector space or a vector space over $\mathbb{Z}_p$, respectively.
  More generally, the scalar can be chosen from any number system in which, roughly speaking, we can add, subtract, multiply, and divide according to the usual laws of arithmetic. In abstract algebra, such a number system is called a field.

- The definition of a vector space does not specify what the set V consist of. Neither does it specify what the operations called "addition" and "scalar multiplication" look like.
- The vector space is defined not only by the set V but also by the operation of addition, multiplication by scalar and the scalar field.

Examples:

- For any $n \geq 1$, $\mathbb{R}^n$ is a vector space with the usual operations of additions and scalar multiplication.
- The set of all 2×3 matrices is a vector space with the usual operations of matrix addition and matrix scalar multiplication.
- The set of all $m \times n$ matrices M_{mn} is a vector space with the usual operations of matrix addition and matrix scalar multiplication.
- Let $\mathcal{P}_2$ denote the set of all polynomials of degree 2 or less with real coefficients: $p(x) \in \mathcal{P}_2$ with $p(x) = a_0 + a_1 x + a_2 x^2$.
- In general, for any $n \geq 0$, the set $\mathcal{P}_n$ of all polynomials of degree less than of equal to n, with real coefficients, is a vector space.
- The set $\mathcal{P}$ of all polynomials, with real coefficients, is a vector space.
- The set $\mathcal{F}$ denote the set of all real-valued functions defined on the real line. If f and g are two such functions and c is a scalar, then $f+g$ and cf are defined by

  $$(f+g)(x) = f(x) + g(x) \tag{1}$$

  and

  $$(cf)(x) = cf(x) \tag{2}$$

  is a vector space.
- The set $\mathcal{F}[a,b]$ denote the set of all real-valued functions defined on some closed interval $[a,b]$ of the real line is a vector space.
- The set $\mathbb{Z}$ of integers with the usual operations is not a vector space since it does not satisfied the axiom the closure under scalar multiplication, for example: for $2 \in \mathbb{Z}$, $\frac{1}{3}(2) \notin \mathbb{Z}$.
- Let $V = \mathbb{R}^2$ with the usual definition of addition but the following definition of scalar multiplication:

  $$c \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} cx \\ 0 \end{bmatrix} \tag{3}$$

  is not a vector space since it fails the axiom 10,

  $$1 \begin{bmatrix} x \\ y \end{bmatrix} \neq \begin{bmatrix} x \\ y \end{bmatrix} \tag{4}$$

Theorem 6.1: Let V be a vector space, $\bar{u}$ a vector in V, and c a scalar.

a. $0\bar{u} = \bar{0}$

b. $c\bar{0} = \bar{0}$

c. $(-1)\bar{u} = -\bar{u}$

d. If $c\bar{u} = \bar{0}$, then $c = 0$ or $\bar{u} = 0$

Proof: See book, pag. 437.

Subtraction: We will write $\bar{u} - \bar{v}$ for $\bar{u} + (-\bar{v})$, defining subtraction of vectors.

Linear combination: We will exploit the associativity property of addition to unambiguously write $\bar{u} + \bar{v} + \bar{w}$ for the sum of three vectors and, more generally, $c_1\bar{u}_1 + c_2\bar{u}_2 + \cdots + c_n\bar{u}_n$ for a linear combination of vectors.

Subspace: A subset W of a vector space V is called a subspace of V if W is itself a vector space with the same scalars, addition, and scalar multiplication as V.

As in $\mathbb{R}^n$, checking to see whether a subset W of a vector space V is a subspace of V involves testing only two of the ten vector space axioms.

Theorem 6.2: Let V be a vector space and let W be a nonempty subset of V. Then W is a subspace of V if and only if the following conditions hold:

a. If $\bar{u}$ and $\bar{v}$ are in W, then $\bar{u} + \bar{v}$ is in W.

b. If $\bar{u}$ is in W and c is a scalar, then $c\bar{u}$ is in W

Proof: See book, pag. 438.

Examples:

a. The set $\mathcal{P}_n$ of all polynomials with degree at most n is a subspace of the vector space $\mathcal{P}$ of all polynomials.

b. The symmetric $n \times n$ matrices form a subspace W of M_{nn}, where M_{nn} is the set of all $n \times n$ matrices with the usual operations of matrix addition and matrix scalar multiplication: Let A and B be in W and let c be a scalar, then (a) $(A+B)^T = A^T + B^T = A + B$ and (b) $(cA)^T = cA^T = cA$.

c. Let $\mathcal{C}$ be the set of all continuous real-valued functions defined on $\mathbb{R}$ and let $\mathcal{D}$ be the set of all differentiable real-valued functions defined on $\mathbb{R}$. It happens that $\mathcal{C}$ and $\mathcal{D}$ are subspaces of $\mathcal{F}$, the vector space of all real-valued functions defined on $\mathbb{R}$:
From calculus, if f and g are continuous functions and c is a scalar, then $f + g$ and cf are also continuous. Hence, $\mathcal{C}$ is closed under addition and scalar multiplication and so is a subspace of $\mathcal{F}$. If f and g are differentiable, then so are $f + g$ and cf, $(f+g)' = f' + g'$ and $(cf)' = c(f')$. So $\mathcal{D}$ is also closed under addition and scalar multiplication, making it a subspace of $\mathcal{F}$.

Comment: It is a theorem of calculus that every differentiable function is continuous. Consequently, $\mathcal{D}$ is contained in $\mathcal{C}$, i.e. $\mathcal{D} \subset \mathcal{C}$, making $\mathcal{D}$ a subspace of $\mathcal{C}$. It is also the case that every polynomial function is differentiable, so $\mathcal{P} \subset \mathcal{D}$, and thus $\mathcal{P}$ a subspace of $\mathcal{D}$. Then we have the following hierarchy of subspaces of $\mathcal{F}$: $\mathcal{P} \subset \mathcal{D} \subset \mathcal{C} \subset \mathcal{F}$.

d. If V is a vector space, then V is clearly a subspace of itself. The set $\{\bar{0}\}$, consisting of only the zero vector, is also a subspace of V, called the zero subspace. This two subspaces are called trivial subspaces of V.

c. If W is a subspace of a vector space V, then (from Theorem 6.2) W contains the zero vector $\{\bar{0}\}$ of V.

Comment to (c): The above observation is consistent with the fact that lines and planes are subspaces of $\mathbb{R}^3$ if and only if the contain the origin. Then, the requirement that every subspace must contain $\{\bar{0}\}$ is sometimes useful in showing that a set is not a subspace.

Exercise for the student in class (Example 6.12): Let $\mathcal{S}$ be the set of all functions that satisfy the differential equation $f'' + f = 0$. Show that $\mathcal{S}$ is a subspace of $\mathcal{F}$. This is an example of homogeneous linear differential equation (the solution sets of such equations are always subspaces of $\mathcal{F}$).
Solution: It is nonempty, since the nil function satisfied the above differential equation. We have to demonstrate that $(c_1 f_1 + c_2 f_2)'' + (c_1 f_1 + c_2 f_2) = 0$.

Exercise for the student in class (Example 6.15): Using the information in 'Comment to (c)', check if W is a subspace of M_{22}, where W is the set of all 2×2 matrices of the form

$$\begin{bmatrix} a & a+1 \\ 0 & b \end{bmatrix} \tag{5}$$

Exercise for the student in class (Example 6.16): Let W be the set of all 2×2 matrices with determinant 0. Demonstrate that W is not a subspace of M_{22}. Help: used for testing the properties the matrices

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{6}$$

Spanning Sets

If $S = \{\bar{v}_1, \bar{v}_2, \cdots, \bar{v}_k\}$ is a set of vectors in a vector space V, then the set of all LC of $\bar{v}_1, \bar{v}_2, \cdots, \bar{v}_k$ is called the **span** of $\bar{v}_1, \bar{v}_2, \cdots, \bar{v}_k$ and is denoted by $span(\bar{v}_1, \bar{v}_2, \cdots, \bar{v}_k)$ or $span(S)$. If $V = span(S)$, then S is called a **spanning set** for V and V is said to be **spanned** by S.

Exercise for the student in class (Example 6.18): Show that
$M_{23} = span(E_{11}, E_{12}, E_{13}, E_{21}, E_{22}, E_{23})$, where

$$E_{11} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad E_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad E_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{7}$$

$$E_{21} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad E_{22} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad E_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{8}$$

Help: you have to show that any 2×3 can be written as LC of the above six matrices.

Exercise for the student in class (Example 6.21):

a. In M_{22}, found the span of the matrices A, B, and C (defined below)

b. Let X an arbitrary symmetric matrix (defined below), write it in term of the above matrices.

$$A = \begin{bmatrix}1 & 1\\1 & 0\end{bmatrix} \quad B = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \quad C = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix} \tag{9}$$

$$X = \begin{bmatrix}x & y\\y & z\end{bmatrix} \tag{10}$$

Solution: [a] you should find that these matrices span the subspace of symmetric matrices. [b] you should find that the coefficients in the expansion $aA + bB + cC$ are $a = x - z$, $b = z$ and $c = -x + y + z$.

Theorem 6.3: Let $\bar{v}_1, \cdots, \bar{v}_k$ be vectors in a vector space V.

a. $span(\bar{v}_1, \cdots, \bar{v}_k)$ is subspace of V

b. $span(\bar{v}_1, \cdots, \bar{v}_k)$ is the smallest subspace of V that contains $\bar{v}_1, \cdots, \bar{v}_k$.

Proof: See book, pag. 445.

Linear Independence (LI), Basis, and Dimension

This section extends the notion of LI, basis, and dimension to general vector spaces. In most cases, the proof of the theorems carry over in Sections 2.3 and 3.5 of the book, can be used replacing $\mathbb{R}^n$ by the vector space V.

Linear Independence (in finite spaces)

A set of vectors $\bar{v}_1, \cdots, \bar{v}_k$ in a vector space V is **linearly dependent** (LD) if there are scalars $c_1, \cdots, c_k$, at least one of which is not zero, such that

$$c_1\,\bar{v}_1 + \cdots + c_k\,\bar{v}_k = \bar{0} \tag{11}$$

A set of vectors that is not LD is said to be **linearly independent.**

Comment: As in $\mathbb{R}^n$, $\{\bar{v}_1, \cdots, \bar{v}_k\}$ is LI in a vector space V if and only if

$$c_1\,\bar{v}_1 + \cdots + c_k\,\bar{v}_k 0 = \bar{0} \Rightarrow c_i = 0 \tag{12}$$

for $i = 1, \cdots, k$.

Example 6.25: Show that the set $\{1, x, x^2, \cdots, x^n\}$ is LI in $\mathcal{P}_n$:
Starting with

$$\sum_{i=0}^{n} c_i x^i = c_0 + c_1 x + \sum_{i=2}^{n} c_i x^i = 0 \tag{13}$$

we set $x = 0$ which gives $c_0 = 0$. Then we take the derivative and get

$$\sum_{i=1}^{n} c_i i\, x^{i-1} = c_1 + c_2 x + \sum_{i=3}^{n} c_i i\, x^{i-1} = 0 \tag{14}$$

and setting again, $x = 0$, we get $c_1 = 0$. By continuing this process we demonstrate that all $c_i = 0$ and then the set $\{1, x, x^2, \cdots, x^n\}$ is LI.

Theorem 6.4 (alternative formulation of LD): A set of vectors $\{\bar{v}_1, \cdots, \bar{v}_k\}$ in a vector space V is LD if and only if at least one of the vectors can be expressed as LC of the others.
Proof: The proof is identical to that of Theorem 2.5.

Example 6.22: In $\mathcal{P}_2$, the set $\{f_1(x), f_2(x), f_3(x)\}$ where

$$f_1(x) = 1 + x + x^2 \quad (15)$$
$$f_2(x) = 1 - x + 3x^2 \quad (16)$$
$$f_3(x) = 1 + 3x - x^2 \quad (17)$$

is LD, since $f_3(x) = 2f_1(x) - f_2(x)$

Exercise for the student in class (Example 6.26): In $\mathcal{P}_2$, determine whether the set $\{f_1(x), f_2(x), f_3(x)\}$ is LI, where

$$f_1(x) = 1 + x \quad (18)$$
$$f_2(x) = x + x^2 \quad (19)$$
$$f_3(x) = 1 + x^2 \quad (20)$$

Help: write $c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) = 0$ and search for the coefficients. You should get that is $\{f_1(x), f_2(x), f_3(x)\}$ is LI.

Linear Independence (in infinite spaces)

A set S of vectors in a vector space V is **linearly dependent** if it contains finitely many LD vectors.
A set of vectors that is not LD is said to be **linearly independent.**

Example 6.28: In $\mathcal{P}$, shows that $S = \{1, x, x^2, \cdots\}$ is LI:
Suppose there is a finite subset T of S that is LD. Let x^m be the highest power of x in T and let x^n be the lowest power of x in T. Then there are scalars $c_n, c_{n+1}, \cdots, c_m$, not all zero, such that

$$\sum_{i=n}^{m} c_i\, x^i = 0 \quad (21)$$

By taking the derivatives we can show that all $c_i = 0$ with $i = n, \cdots, m$, which is a contradiction. Hence, S cannot contain finitely many linearly dependent vectors, so it is LI.

Bases

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

1. $\mathcal{B}$ spans V and
2. $\mathcal{B}$ is linearly independent

Examples:

- If $\bar{e}_i$ is the ith column of the $n \times n$ matrix, the $\{\bar{e}_1, \bar{e}_2, \cdots, \bar{e}_n\}$ is a basis for $\mathbb{R}^n$, called the **standard basis** for $\mathbb{R}^n$.
- $\{1, x, x^2, \cdots, x^n\}$ is a basis for $\mathcal{P}_n$, called the **standard basis** for $\mathcal{P}_n$.
- The set $\mathcal{E} = \{E_{11}, \cdots, E_{1n}, E_{21}, \cdots, E_{2n}, E_{m1}, \cdots, E_{mn}\}$ is a basis for M_{mn} where the matrices E_{ij} are as defined above. $\mathcal{E}$ is called the **standard basis** for M_{mn}.

Example 6.32: Show that $\mathcal{B} = \{1 + x, x + x^2, 1 + x^2\}$ is a basis for $\mathcal{P}_2$.
Solution: It was already shown that $\mathcal{B}$ is LI. Next we must show that there are scalars c_1, c_2, and c_3 such that $a + bx + cx^2 = c_1(1 + x) + c_2(x + x^2) + c_3(1 + x^2)$. We built a linear system and show that the range of the matrix is 3 (not need to show only that the solution exist).

Coordinates

The most useful aspect of coordinate vectors is that they allow us to transfer information from a general vector space to $\mathbb{R}^n$, where we have the tools defines in $\mathbb{R}^n$ at our disposal.

Theorem 6.5: Let V be a vector space and let $\mathcal{B}$ be a basis for V. For every vector $\bar{v}$ in V, there is exactly one way to write $\bar{v}$ as a LC of the basis vectors in $\mathcal{B}$.
Proof: It is the same as the one of Theorem 3.29.

Unique representation: The converse of Theorem 6.5 is also true. That is, if $\mathcal{B}$ is a set of vectors in a vector space V with the property that every vector in V can be written uniquely as a LC of the vectors in $\mathcal{B}$, then $\mathcal{B}$ is a basis for V. In this sense, the **unique representation property** characterizes a basis.

Coordinates: Let $\mathcal{B} = \{\bar{v}_1, \cdots, \bar{v}_n\}$ be a basis for a vector space V. Let $\bar{v}$ be a vector in V, and write $\bar{v} = c_1\bar{v}_1 + \cdots + c_n\bar{v}_n$. Then $c_1, \cdots, c_n$ are called the **coordinates of $\bar{v}$ with respect to $\mathcal{B}$**, and the column vector

$$[\bar{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \tag{22}$$

is called the **coordinates vector of $\bar{v}$ with respect to $\mathcal{B}$**.

Example: The coordinate vector of a polynomial $p(x) = a_0 + a_1x + \cdots + a_nx^n$ in $\mathcal{P}_n$ with respect to the standard basis $\mathcal{B} = \{1, x, \cdots, x^n\}$ is the vector

$$[p(x)]_{\mathcal{B}} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} \tag{23}$$

in $\mathbb{R}^{n+1}$.

Example 6.35: Find the coordinate vectors $[p(x)]_{\mathcal{B}_i}$ of $p(x) = 2 - 3x + 5x^2$ with respect to the standard bases $\mathcal{B}_1 = \{1, x, x^2\}$ and $\mathcal{B}_2 = \{x^2, x, 1\}$.

$$[p(x)]_{\mathcal{B}_1} = \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix} \tag{24}$$

$$[p(x)]_{\mathcal{B}_2} = \begin{bmatrix} 5 \\ -3 \\ 2 \end{bmatrix} \tag{25}$$

Exercise for the student in class (Example 6.36): Find the coordinates vector $[A]_\mathcal{B}$ of

$$A = \begin{bmatrix} 2 & -1 \\ 4 & 3 \end{bmatrix} \tag{26}$$

with respect to the standard basis $\mathcal{B} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ of M_{22}.
Solution: you should get:

$$[A]_\mathcal{B} = \begin{bmatrix} 2 \\ -1 \\ 4 \\ 3 \end{bmatrix} \tag{27}$$

Theorem 6.6: Let $\mathcal{B} = \{\bar{v}_1, \cdots, \bar{v}_n\}$ be a basis for a vector space V. Let $\bar{u}$ and $\bar{v}$ be vectors in V and let c be scalar. Then

a. $[\bar{u} + \bar{v}]_\mathcal{B} = [\bar{u}]_\mathcal{B} + [\bar{v}]_\mathcal{B}$

b. $[c\bar{u}]_\mathcal{B} = c[\bar{u}]_\mathcal{B}$

Proff: See book, pag. 455.

Corollary to the Theorem 6.6: Coordinate vectors preserves lineal combinations:

$$[c_1\bar{u}_1 + \cdots + c_k\bar{u}_k]_\mathcal{B} = c_1[\bar{u}_1]_\mathcal{B} + \cdots + c_k[\bar{u}_k]_\mathcal{B} \tag{28}$$

Theorem 6.7: Let $\mathcal{B} = \{\bar{v}_1, \cdots, \bar{v}_n\}$ be a basis for a vector space V and let $\{\bar{u}_1, \cdots, \bar{u}_k\}$ be vectors in V. Then $\{\bar{u}_1, \cdots, \bar{u}_k\}$ is LI in V if and only if $\{[\bar{u}_1]_\mathcal{B}, \cdots, [\bar{u}_k]_\mathcal{B}\}$ is LI in $\mathbb{R}^n$.
Prof $\Rightarrow$: Assume that $\{\bar{u}_1, \cdots, \bar{u}_k\}$ is LI in V and let

$$c_1[\bar{u}_1]_\mathcal{B} + \cdots + c_k[\bar{u}_k]_\mathcal{B} = \bar{0} \tag{29}$$

in $\mathbb{R}^n$, with unknown c_i.

Using the property from the Corollary above we have

$$[c_1\bar{u}_1 + \cdots + c_k\bar{u}_k]_\mathcal{B} = \bar{0} \tag{30}$$

this means that the coordinates of the vector $c_1\bar{u}_1 + \cdots + c_k\bar{u}_k$ with respect to $\mathcal{B}$ are all zero. Explicitly this implies

$$c_1\bar{u}_1 + \cdots + c_k\bar{u}_k = 0\bar{v}_1 + \cdots + 0\bar{v}_n = \bar{0} \tag{31}$$

Since we assume that $\{\bar{u}_1, \cdots, \bar{u}_k\}$ is LI, the above relation implies that $c_i = 0$ for $i = 1, \cdots, k$, so $\{[\bar{u}_1]_\mathcal{B}, \cdots, [\bar{u}_k]_\mathcal{B}\}$ is LI in $\mathbb{R}^n$.

The converse implication is left for the student (Exercise 32 of the book). Help: applies similar ideas used above.

Comment: Observe that, in the special case where $\bar{u}_i = \bar{v}_i$ we have

$$\bar{v}_i = 0\bar{v}_1 + \cdots + 1\,\bar{v}_i + \cdots + 0\bar{v}_n = \bar{0} \tag{32}$$

so

$$[\bar{v}_i]_\mathcal{B} = \bar{e}_i \tag{33}$$

and

$$\{[\bar{v}_1]_\mathcal{B}, \cdots, [\bar{v}_n]_\mathcal{B}\} = \{[\bar{e}_1], \cdots, [\bar{e}_n]\} \tag{34}$$

Dimension:

Dimension is the number of vectors in a basis. Since a vector space can have more that one basis, we need to show that this definition makes sense.

Theorem 6.8: Let $\mathcal{B} = \{\bar{v}_1, \cdots, \bar{v}_n\}$ be a basis for a vector space V.

a. Any set of more than n vectors in V must be LD

b. Any set of fewer than n vectors in V cannot span V.

Proof: See book, pag. 456.

The Basis Theorem (Theorem 6.9): If a vector space V has a basis with n vectors, then every basis for V has exactly n vectors.
Proof: The proof of Theorem 3.23 also works here. See book, pag. 457.

Definitions:

Finite-dimensional: A vector space V is called finite-dimensional if it has a basis consisting of finitely many vectors.

Dimension: The dimension of V, denoted by $\dim V$, is the number of vectors in a basis for V. The dimension of the zero vector space $\{\bar{0}\}$ is defined to be zero.

Infinite-dimensional: A vector space that has no finite basis is called infinite-dimensional.

Example 6.38: Since the standard basis for $\mathbb{R}^n$ has n vectors, $\dim\mathbb{R}^n = n$. In the case of $\mathbb{R}^3$:

- a one-dimensional subspace is span by a non-zero vector, i.e. a line through the origin,
- a two-dimensional subspace is spanned by its basis of two LI vectors, i.e. a plane through the origin
- any three LI vectors must span $\mathbb{R}^3$.

The subspaces of $\mathbb{R}^3$ are now completely classified according to dimension (see table 1):

$\dim V$	V
3	$\mathbb{R}^3$
2	Plane through $\bar{0}$
1	Line through $\bar{0}$
0	$\{\bar{0}\}$

Table 1: Dimensions.

Examples:

- The standard basis for $\mathcal{P}_n$ contains $n+1$ vectors, so $\dim\mathcal{P}_n = n+1$
- The standard basis for M_{mn} contains mn (léase: producto de m por n) vectors, so $\dim M_{mn} = mn$.
- Both $\mathcal{P}$ and $\mathcal{F}$ are infinite-dimensional, since they each contain the infinite LI set $\{1, x, x^2, \cdots\}$.

Exercise (Example 6.42): Find the dimension of the vector space W of symmetric 2×2 matrices.
Solution: A symmetric 2×2 matrix has the form,

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} = a_{11} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_{12} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + a_{22} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{35}$$

so W is expanded by the set

$$\mathcal{S} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\} \tag{36}$$

If $\mathcal{S}$ is LI, then it will be a basis for W. Setting

$$a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{37}$$

we get

$$a = b = c = 0 \tag{38}$$

Hence $\mathcal{S}$ is LI and dim$W$=3.

Comment: The dimension of a vector space V provides with much information about V and can greatly simplify the work needed in certain types of calculations, as the next few theorems and examples illustrate.

Theorem 6.10: Let V be a vector space with dim$V$=n. Then

a. Any LI set in V contains at most n vectors.

b. Any spanning set for V contains at least n vectors.

c. Any LI set of exactly n vectors in V is a basis for V.

d. Any spanning set for V consisting of exactly n vectors is a basis for V.

e. Any LI set in V can be extended to a basis for V.

f. Any spanning set for V can be reduced to a basis for V.

Proof: See book, pag. 459.

Example 6.43: Let us see whether S is a basis for V for the following cases

a. $V = \mathcal{P}_2$ and $S = \{1 + x, 2 - x + x^2, 3x - 2x^2, -1 + 3x + x^2\}$.
Solution: since dim$\mathcal{P} = 2 + 1 = 3$ and S contains 4 vectors, S is LD, hence S is not, by Theorem 6.10(a), a basis for $\mathcal{P}_2$.

b. $V = M_{22}$, $\mathcal{S} = \left\{ \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \right\}$.
Solution: since dim$M_{22} = 2 \times 2 = 4$ and S contains three vectors, S cannot, by Theorem 6.10(b), span M_{22}. Hence, S is not a basis for M_{22}.

c. $V = \mathcal{P}_2$ and $S = \{1 + x, x + x^2, 1 + x^2\}$.
Solution: since dim$\mathcal{P} = 2 + 1 = 3$ and S contains 3 vectors, S will be a basis is they are LI or if it spans $\mathcal{P}_2$, by Theorem 6.10(c) or (d). It was demonstrated earlier that they are LI, **do it again!**. Therefore, S is a basis for $\mathcal{P}_2$.

Theorem 6.11: Let W be a subspace of a finite-dimensional vector space V. Then

a. W is finite-dimensional and dim$W \leq$dim$V$.

b. dim$W$=dim$V$ if and only if $W = V$.

Proof: See book, pag. 460.

Change of Basis

In many applications, a problem described using one coordinate system may be solved more easily by switching to a new coordinate system.

Introduction

From Theorem 3.29: Let S be a subspace of $\mathbb{R}^n$ and let $\mathcal{B} = \{\bar{v}_1, \cdots, \bar{v}_k\}$ be a basis for S. For every vector $\bar{v}$ in S, there is exactly one way to write $\bar{v}$ as a linear combination of the basis vectors in $\mathcal{B}$:

$$\bar{v} = c_1\bar{v}_1 + \cdots + c_k\bar{v}_k \tag{39}$$

After this Theorem we defined coordinates: Let S be a subspace of $\mathbb{R}^n$ and let $\mathcal{B} = \{\bar{v}_1, \cdots, \bar{v}_k\}$ be a basis for S. Let $\bar{v}$ be a vector in S, and write $\bar{v} = c_1\bar{v}_1 + \cdots + c_k\bar{v}_k$. Then $c_1, \cdots, c_k$ are called the coordinates of $\bar{v}$ with respect to $\mathcal{B}$, and the column vector

$$[\bar{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix} \tag{40}$$

is called the coordinate vector of $\bar{v}$ with respect to $\mathcal{B}$.

Consider the following two different coordinate systems for $\mathbb{R}^2$,

- $\mathcal{B} = \{\bar{u}_1, \bar{u}_2\}$ $\bar{u}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\bar{u}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$
- $\mathcal{C} = \{\bar{v}_1, \bar{v}_2\}$ $\bar{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\bar{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

The vector $\bar{r} = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$, can be written in terms of any of these two basis:

$$\bar{r} = \bar{u}_1 + 3\bar{u}_2 \tag{41}$$
$$\bar{r} = 6\bar{v}_1 - \bar{v}_2 \tag{42}$$

The coordinates of the vector $\bar{r}$ with respect to $\mathcal{B}$ and $\mathcal{C}$ are

$$[\bar{r}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \tag{43}$$

$$[\bar{r}]_{\mathcal{C}} = \begin{bmatrix} 6 \\ -1 \end{bmatrix} \tag{44}$$

We can find the coordinates of one basis in terms of the other with the following procedures (see also Example 6.45). In this case we calculate the coordinate in the basis $\mathcal{C}$ from the coordinates in the basis $\mathcal{B}$ by expanding the basis vectors $\bar{u}_i$ in the basis $\mathcal{C}$,

$$\begin{aligned}
\bar{r} &= \begin{bmatrix} 5 \\ -1 \end{bmatrix} = \bar{u}_1 + 3\bar{u}_2 && (45)\\
&= \left\{-3\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\} + 3\left\{3\begin{bmatrix} 1 \\ 0 \end{bmatrix} - 1\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\} && (46)\\
&= \{-3\bar{v}_1 + 2\bar{v}_2\} + 3\{3\bar{v}_1 - 1\bar{v}_2\} && (47)\\
&= 6\bar{v}_1 - \bar{v}_2 && (48)
\end{aligned}$$

Hence, we recover the previous result

$$[\bar{r}]_{\mathcal{C}} = \begin{bmatrix} 6 \\ -1 \end{bmatrix} \quad (49)$$

This procedure gives as a systematic way to transform the vector coordinates for a given vector from one basis to another by using a matrix transformation (see next section).

Change-of-Basis Matrices

Let us systematized the above procedure. From $\bar{r} = \bar{u}_1 + 3\bar{u}_2$, we have

$$[\bar{r}]_{\mathcal{C}} = [\bar{u}_1 + 3\bar{u}_2]_{\mathcal{C}} = [\bar{u}_1]_{\mathcal{C}} + 3\,[\bar{u}_2]_{\mathcal{C}} \quad (50)$$

by Theorem 6.6.

Thus,

$$[\bar{r}]_{\mathcal{C}} = [[\bar{u}_1]_{\mathcal{C}}\,[\bar{u}_2]_{\mathcal{C}}]\begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad (51)$$

The coordinates of $\bar{u}_i$ in the basis $\mathcal{C}$ are,

$$\bar{u}_1 = -3\bar{v}_1 + 2\bar{v}_2 \Rightarrow [\bar{u}_1]_{\mathcal{C}} = \begin{bmatrix} -3 \\ 2 \end{bmatrix} \quad (52)$$

$$\bar{u}_2 = 3\bar{v}_1 - \bar{v}_2 \Rightarrow [\bar{u}_2]_{\mathcal{C}} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \quad (53)$$

Then,

$$\begin{aligned}
[\bar{r}]_{\mathcal{C}} &= \begin{bmatrix} -3 & 3 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix} && (54)\\
&= P[\bar{r}_{\mathcal{B}}] && (55)
\end{aligned}$$

where P is the matrix whose columns are $[\bar{u}_1]_{\mathcal{C}}$ and $[\bar{u}_2]_{\mathcal{C}}$.

Luego, si llamamos base vieja a la base $\mathcal{B}$, y base nueva a la base $\mathcal{C}$, la matriz P transforma los coeficientes de un dado vector de la base vieja a la nueva. De este modo si el vector $\bar{r}$ en la base vieja se excribía como $\bar{r} = \bar{u}_1 + 3\bar{u}_2$, en la nueva se escribirá $\bar{r} = c_1\bar{v}_1 + c_2\bar{v}_2$ con

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = [\bar{r}]_{\mathcal{C}} = P[\bar{r}_{\mathcal{B}}] = \begin{bmatrix} -3 & 3 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ -1 \end{bmatrix}$$

esto es, $\bar{r} = 6\bar{v}_1 - \bar{v}_2$.

Change of basis: Let $\mathcal{B} = \{\bar{u}_1, \cdots, \bar{u}_n\}$ and $\mathcal{C} = \{\bar{v}_1, \cdots, \bar{v}_n\}$ be bases for a vector space V. The $n \times n$ matrix whose columns are the coordinate vectors $[\bar{u}_1]_\mathcal{C}, \cdots, [\bar{u}_n]_\mathcal{C}$ of the vectors in $\mathcal{B}$ with respect to $\mathcal{C}$ is denoted by $P_{\mathcal{C}\leftarrow\mathcal{B}}$ and is called the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{C}$. That is,

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = [[\bar{u}_1]_\mathcal{C}\, [\bar{u}_2]_\mathcal{C} \cdots [\bar{u}_n]_\mathcal{C}] \tag{56}$$

Think of $\mathcal{B}$ as the "old" basis and $\mathcal{C}$ as the "new" basis. Then the columns of $P_{\mathcal{C}\leftarrow\mathcal{B}}$ are just the coordinate vectors obtained by writing the old basis vectors in terms of the new ones.

Theorem 6.12: Let $\mathcal{B} = \{\bar{u}_1, \cdots, \bar{u}_n\}$ and $\mathcal{C} = \{\bar{v}_1, \cdots, \bar{v}_n\}$ be bases for a vector space V and let $P_{\mathcal{C}\leftarrow\mathcal{B}}$ be the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{C}$. Then

a. $P_{\mathcal{C}\leftarrow\mathcal{B}}[\bar{x}]_\mathcal{B} = [\bar{x}]_\mathcal{C}$ for all $\bar{x}$ in V.

b. $P_{\mathcal{C}\leftarrow\mathcal{B}}$ is the unique matrix P with the property that $P[\bar{x}]_\mathcal{B} = [\bar{x}]_\mathcal{C}$ for all $\bar{x}$ in V.

c. $P_{\mathcal{C}\leftarrow\mathcal{B}}$ is invertible and $(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} = P_{\mathcal{B}\leftarrow\mathcal{C}}$.

Proof: See book, pag. 469.

Comment: A change of basis is a transformation from $\mathbb{R}^n$ to itself that switches from one coordinate system to another. The transformation $P_{\mathcal{C}\leftarrow\mathcal{B}}$ takes $[\bar{x}]_\mathcal{B}$ as an input and returns $[\bar{x}]_\mathcal{C}$ as output. While $P_{\mathcal{B}\leftarrow\mathcal{C}}$ is the other way around. See Fig. 1

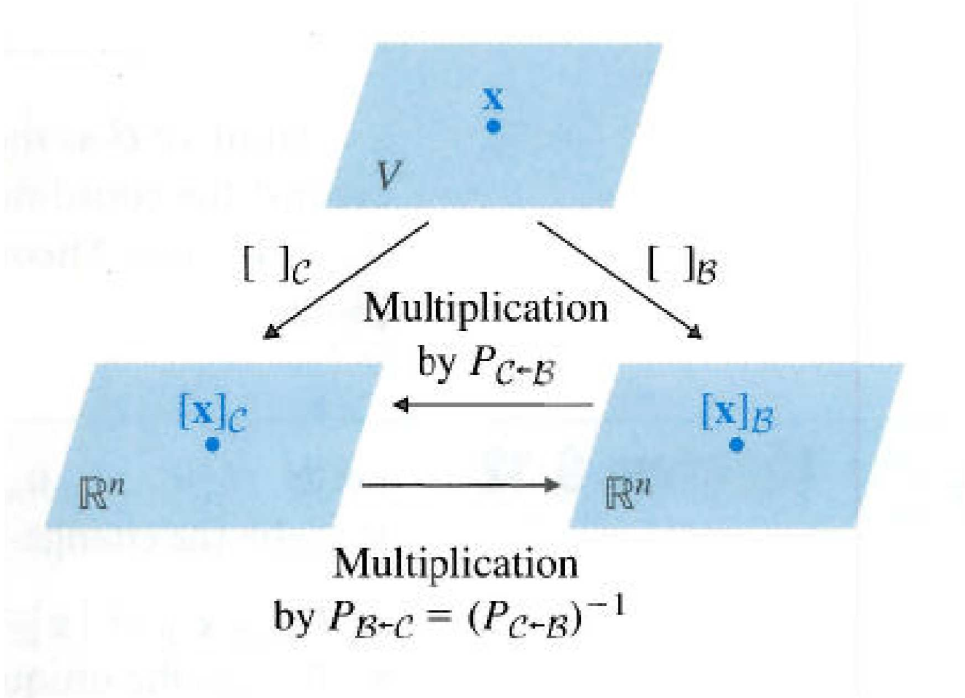

Figure 1: (from book)

Exercise (Example 6.46): Find the change-of-basis matrices $P_{\mathcal{C}\leftarrow\mathcal{B}}$ and $P_{\mathcal{B}\leftarrow\mathcal{C}}$ for the basis $\mathcal{B} = \{1, x, x^2\} = \{\bar{u}_1, \bar{u}_2, \bar{u}_3\}$ and $\mathcal{C} = \{1 + x, x + x^2, 1 + x^2\} = \{\bar{v}_1, \bar{v}_2, \bar{v}_3\}$ of $\mathcal{P}_2$. Then find the coordinate vector of $p(x) = 1 + 2x - x^2$ with respect to $\mathcal{C}$.

Solution: Changing to a standard basis is easy, so we start finding the transformation $P_{\mathcal{B}\leftarrow\mathcal{C}}$. The coordinates vectors for $\mathcal{C}$ in terms of $\mathcal{B}$ are

$$\bar{v}_1 = 1 + x = \bar{u}_1 + \bar{u}_2 \Rightarrow [1+x]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \tag{57}$$

$$\bar{v}_2 = x + x^2 = \bar{u}_2 + \bar{u}_3 \Rightarrow \left[x+x^2\right]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \tag{58}$$

$$\bar{v}_3 = 1 + x^2 = \bar{u}_1 + \bar{u}_3 \Rightarrow \left[x+x^2\right]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \tag{59}$$

Then

$$P_{\mathcal{B}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \tag{60}$$

By direct calculation of the inverse of $P_{\mathcal{B}\leftarrow\mathcal{C}}$ or by expressing the vectors in $\mathcal{B}$ as a linear combination of vectors in $\mathcal{C}$, we should get (**Try this last option at home**)

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 1/2 & 1/2 & -1/2 \\ -1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \end{bmatrix} \tag{61}$$

The coordinate vector of $p(x) = 1 + 2x - x^2$ with respect to $\mathcal{C}$ can be obtained by observing that the coordinate vector of $p(x)$ with respect to $\mathcal{B}$ is

$$[p(x)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \tag{62}$$

and then change from one basis to the other using $P_{\mathcal{C}\leftarrow\mathcal{B}}$,

$$[p(x)]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}}[p(x)]_{\mathcal{B}} \tag{63}$$

$$= \begin{bmatrix} 1/2 & 1/2 & -1/2 \\ -1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \tag{64}$$

$$= \begin{bmatrix} 1/2 + 1 + 1/2 \\ -1/2 + 1 - 1/2 \\ 1/2 - 1 - 1/2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} \tag{65}$$

Exercise for the student in class (Example 6.47): In M_{22}, let $\mathcal{B}$ be the basis $\{E_{11}, E_{21}, E_{12}, E_{22}\}$ with

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad E_{21} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

and let be $\mathcal{C}$ be the basis $\{A, B, C, D\}$ with

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{66}$$

Find the change-of-basis matrix $P_{\mathcal{C}\leftarrow\mathcal{B}}$ and verify that $[X]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}}[X]_{\mathcal{B}}$ for $X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

Solution: We need to express the matrices E_{ij} in the basis $\mathcal{C}$;

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = A \Rightarrow [E_{11}]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{67}$$

$$E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = C - B = -B + C \Rightarrow [E_{21}]_{\mathcal{C}} = \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}$$

$$E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = B - A = -A + B \Rightarrow [E_{12}]_{\mathcal{C}} = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = D - C = -C + D \Rightarrow [E_{12}]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

The matrix which changes from basis $\mathcal{B}$ to $\mathcal{C}$ is

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{68}$$

Finally, verify that $[X]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}}[X]_{\mathcal{B}}$ for $X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$
The coordinate vector of X in the basis $\mathcal{B}$ is

$$X = 1E_{10} + 3E_{21} + 2E_{12} + 4E_{22} \Rightarrow [X]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix} \tag{69}$$

Then,

$$P_{\mathcal{C}\leftarrow\mathcal{B}}[X]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix} \tag{70}$$

$$= \begin{bmatrix} 1+0-2+0 \\ 0-3+2+0 \\ 0+3+0-4 \\ 0+0+0+4 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 4 \end{bmatrix} \tag{71}$$

This implies that X expanded in the basis $\mathcal{C}$ should be given by

$$X = -1A - 1B - 1C + 4D \tag{72}$$

Let us check this,

$$\begin{aligned} -1A - 1B - 1C + 4D &= (-1)\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + (-1)\begin{bmatrix}1 & 1\\0 & 0\end{bmatrix} + (-1)\begin{bmatrix}1 & 1\\1 & 0\end{bmatrix} + 4\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix} \\ &= \begin{bmatrix}-1-1-1+4 & 0-1-1+4\\ 0+0-1+4 & 0+0+0+4\end{bmatrix} \\ &= \begin{bmatrix}1 & 2\\3 & 4\end{bmatrix} = X \end{aligned} \tag{73}$$

Then, it is true.

The Gauss-Jordan Method for Computing a Change-of-Basis Matrix

Finding the change-of-basis matrix to a standard basis is easy and can be done by inspection. Finding the change-of-basis matrix from a standard basis is almost as easy, but requires the calculation of a matrix inverse. From the example 6.46 we have that we can find $[p(x)]_{\mathcal{C}}$ from $[p(x)]_{\mathcal{B}}$ and $P_{\mathcal{B}\leftarrow\mathcal{C}}$ using Gaussian elimination, i.e. row reduction

$$[P_{\mathcal{B}\leftarrow\mathcal{C}} \mid [p(x)]_{\mathcal{B}}] \rightarrow \left[I \mid (P_{\mathcal{B}\leftarrow\mathcal{C}})^{-1}[p(x)]_{\mathcal{B}}\right] = [I \mid P_{\mathcal{C}\leftarrow\mathcal{B}}[p(x)]_{\mathcal{B}}] = [I \mid [p(x)]_{\mathcal{C}}]$$

We now look at a modification of the Gauss-Jordan method that can be used to find the change-of-basis matrix between two nonstandard bases, as in Example 6.47.

Suppose $\mathcal{B} = \{\bar{u}_1, \cdots, \bar{u}_n\}$ and $\mathcal{C} = \{\bar{v}_1, \cdots, \bar{v}_n\}$ are bases for a vector space V and $P_{\mathcal{C}\leftarrow\mathcal{B}}$ is the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{C}$. The ith column of P is

$$[\bar{u}_i]_{\mathcal{C}} = \begin{bmatrix} p_{1i} \\ \vdots p_{ni} \end{bmatrix} \tag{74}$$

so $\bar{u}_i = p_{1i}\bar{v}_1 + \cdots + p_{ni}\bar{v}_n$. If $\mathcal{E}$ is any basis for V, then

$$[\bar{u}_i]_{\mathcal{E}} = [p_{1i}\bar{v}_1 + \cdots + p_{ni}\bar{v}_n]_{\mathcal{E}} = p_{1i}[\bar{v}_1]_{\mathcal{E}} + \cdots + p_{ni}[\bar{v}_n]_{\mathcal{E}} \tag{75}$$

In matrix form,

$$[\bar{u}_i]_{\mathcal{E}} = [[\bar{v}_1]_{\mathcal{E}} \cdots [\bar{v}_n]_{\mathcal{E}}] \begin{bmatrix} p_{1i} \\ \vdots \\ p_{ni} \end{bmatrix} \tag{76}$$

which we can solve by applying Gauss-Jordan elimination to the augmented matrix

$$[[\bar{v}_1]_{\mathcal{E}} \cdots [\bar{v}_n]_{\mathcal{E}} \mid [\bar{u}_i]_{\mathcal{E}}] \tag{77}$$

There are n such systems of equations to be solved, one for each column of $P_{\mathcal{C}\leftarrow\mathcal{B}}$ but the coefficients matrix $[[\bar{v}_1]_{\mathcal{E}} \cdots [\bar{v}_n]_{\mathcal{E}}]$ is the same in each case. Hence, we can solve all the systems simultaneously by row reducing the $n \times 2n$ augmented matrix

$$[[\bar{v}_1]_{\mathcal{E}} \cdots [\bar{v}_n]_{\mathcal{E}} \mid [\bar{u}_1]_{\mathcal{E}}] \cdots [\bar{u}_n]_{\mathcal{E}}] = [C \mid B] \tag{78}$$

Since $\{\bar{v}_1, \cdots, \bar{v}_n\}$ is LI, so is $\{[\bar{v}_1]_{\mathcal{E}} \cdots [\bar{v}_n]_{\mathcal{E}}\}$, by Theorem 6.7. Therefore, the matrix C whose columns are $[\bar{v}_1]_{\mathcal{E}} \cdots [\bar{v}_n]_{\mathcal{E}}$ has the $n \times n$ identity matrix I for its reduced row echelon form, by the Fundamental Theorem. If follows that Gauss-Jordan elimination will necessarily produce

$$[C \mid B] \rightarrow [I \mid P] \tag{79}$$

where

$$P = P_{\mathcal{C}\leftarrow\mathcal{B}} \tag{80}$$

The above prove the following theorem.

Theorem 6.13: $\mathcal{B} = \{\bar{u}_1, \cdots, \bar{u}_n\}$ and $\mathcal{C} = \{\bar{v}_1, \cdots, \bar{v}_n\}$ be bases for a vector space V. Let $B = [[\bar{u}_1]_\mathcal{E} \cdots [\bar{u}_n]_\mathcal{E}]$ and $C = [[\bar{v}_1]_\mathcal{E} \cdots [\bar{v}_n]_\mathcal{E}]$, where $\mathcal{E}$ is any basis for V. Then row reduction applied to the $n \times 2n$ augmented matrix $[C|B]$ produces

$$[C \,|\, B] \to [I \,|\, P_{\mathcal{C}\leftarrow\mathcal{B}}] \tag{81}$$

If $\mathcal{E}$ is a standard basis, this method is particularly easy to use, since in that case

$$B = P_{\mathcal{E}\leftarrow\mathcal{B}} \tag{82}$$

and

$$C = P_{\mathcal{E}\leftarrow\mathcal{C}} \tag{83}$$

Example 6.48 Rework Example 6.47 using Gauss-Jordan method.
Solution:

$$B = P_{\mathcal{E}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{84}$$

and

$$C = P_{\mathcal{E}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{85}$$

Row reduction produces

$$\begin{aligned} [C \,|\, B] &= \left[\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \\ &\to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] = [I \,|\, P_{\mathcal{C}\leftarrow\mathcal{B}}] \end{aligned} \tag{86}$$

Linear Transformations

In this section we extend the concept of linear transformation between arbitrary vector spaces. Linear transformation was already introduced in section 3.6 in the context of matrix transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$.

Linear Transformation (LT): A LT from a vector space V to a vector space W is a mapping $T : V \to W$ such that, for all $\bar{u}$ and $\bar{v}$ in V and for all scalars c,

1. $T(\bar{u} + \bar{v}) = T(\bar{u}) + T(\bar{v})$
2. $T(c\bar{u}) = cT(\bar{u})$

This definition is equivalent to the requirement that T preserve all linear combination, i.e. $T : V \to W$ is a LT if and only if

$$T(c_1\bar{u}_1 + \cdots + c_n\bar{u}_n) = c_1T(\bar{u}_1) + \cdots + c_nT(\bar{u}_n) \tag{87}$$

for all $\bar{u}_i$ in V and scalars c_i with $i = 1, \cdots, n$.

Examples of LT:

- Every matrix transformation is a LT, i.e., if A is an $m \times n$ matrix, then $T_A : \mathbb{R}^n \to \mathbb{R}^n$ defined by $T_A(\bar{x}) = A\bar{x}$, with $\bar{x}$ in $\mathbb{R}^n$ is a LT. **Show it!!**
- The transformation $T : M_{nn} \to M_{nn}$ defined by $T(A) = A^T$ is a LT. **Show it!!**
- The differential operator $D : \mathcal{D} \to \mathcal{F}$ defined by $D(f) = f'$ is a LT. **Show it!!**
- The integration $S : \mathcal{C}[a, b] \to \mathbb{R}$ defined as $S(f) = \int_a^b f(x)dx$ is a LT. **Show it!!**

Examples of non LT: Show them!!

1. $T : M_{22} \to \mathbb{R}$ defined by $T(A) = detA$
2. $T : \mathbb{R} \to \mathbb{R}$ defined by $T(x) = 2^x$
3. $T : \mathbb{R} \to \mathbb{R}$ defined by $T(x) = x + 1$

Special transformations:

Zero transformation: For any vector spaces V and W, the transformation $T_0 : V \to W$ that maps every vector V to the zero vector in W is called zero transformation: $T_0(\bar{v}) = \bar{0}$ for all $\bar{v}$ in V.

Identity transformation: For any vector space V, the transformation $I : V \to V$ that maps every vector in V to itself is called the identity transformation: $I(\bar{v}) = \bar{v}$ for all $\bar{v}$ in V.

Properties of LT

In chapter 3 of the book, all LT were matrix transformations, and their properties were directly related to properties of the matrices involved. The following Theorem is easy to demonstrate for matrices but it takes a bit more care for the general case.

Theorem 6.14: Let $T : V \to W$ be a LT. Then

a. $T(\bar{0}) = \bar{0}$

b. $T(-\bar{v}) = -T(\bar{v})$ for all $\bar{v}$ in V.

c. $T(\bar{u} - \bar{v}) = T(\bar{u}) - T(\bar{v})$ for all $\bar{u}$ and $\bar{v}$ in V.

Proff: See book, pag. 479.

Comments:

- Property (a) can be useful in showing that certain transformation is not linear. For example $T(x) = 2^x$.
- Be warned, however, that there are lots of transformations that do map the zero vector to the zero vector but that are still **no** linear, for example $T : M_{22} \to \mathbb{R}$ defined by $T(A) = detA$.
- The most important property of a LT $T : V \to W$ is that T is completely determined by its effect on a basis for V. The example below shows what this means.

Example 6.55: Suppose T is a LT from $\mathbb{R}^2$ to $\mathcal{P}_2$ such that

$$\begin{aligned} T\begin{bmatrix}1\\1\end{bmatrix} &= 2-3x+x^2 \\ T\begin{bmatrix}2\\3\end{bmatrix} &= 1-x^2 \end{aligned} \tag{88}$$

Find

$$T\begin{bmatrix}-1\\2\end{bmatrix} \tag{89}$$

Procedure:

1. First we expand each vector in the basis $\mathcal{B}$ form by the two LI vectors

$$\mathcal{B}=\left\{\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}2\\3\end{bmatrix}\right\} \tag{90}$$

2. Then, we find the coefficients of the expansion
3. Finally we use the fact that T is linear.

Solution:

$$\begin{bmatrix}-1\\2\end{bmatrix}=c_1\begin{bmatrix}1\\1\end{bmatrix}+c_2\begin{bmatrix}2\\3\end{bmatrix}\Rightarrow c_1=-7,\quad c_2=3 \tag{91}$$

then

$$\begin{aligned} T\begin{bmatrix}-1\\2\end{bmatrix} &= -7T\begin{bmatrix}1\\1\end{bmatrix}+3T\begin{bmatrix}2\\3\end{bmatrix} \\ &= -7(2-3x+x^2)+3(1-x^2)=-14+21x-7x^2+3-3x^2 \\ &= -11+21x-10x^2 \end{aligned} \tag{92}$$

Theorem 6.15: Let $T:V\to W$ be a LT and let $\mathcal{B}=\{\bar{v}_1,\cdots,\bar{v}_n\}$ be a spanning set for V. Then $T(\mathcal{B})=\{T(\bar{v}_1),\cdots,T(\bar{v}_n)\}$ spans the range of T (see the definition of *range* below).

Starts remembering from chapter 3 of the book: Linear Transformation...

Transformation: More generally, a transformation (or mapping or function) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector $\bar{v}$ in $\mathbb{R}^n$ a unique vector $T\bar{v}$ in $\mathbb{R}^m$.

Domain-Codomain: The domain of T is $\mathbb{R}^n$, and the codomain of T is $\mathbb{R}^m$. This is indicated by writing $T:\mathbb{R}^n\to\mathbb{R}^m$.

Image: For a vector $\bar{v}$ in the domain of T, the vector $T(\bar{v})$ in the codomain is called the image of $\bar{v}$ under (the action of) T.

Range: The set of all possible images $T(\bar{v})$ is called the range of T.

...Ends remenmbering from chapter 3 of the book: Linear Transformation

Proof: The range of T is the set of all vectors in W that are of the form $T(\bar{v})$, where $\bar{v}$ is in V. Let $T(\bar{v})$ be in the range of T. Since $\mathcal{B}$ spans V, there are scalars $c_1, \cdots, c_n$ such that

$$\bar{v} = c_1\bar{v}_1 + \cdots + c_n\bar{v}_n \tag{93}$$

Applying T and using the fact that it is a LT, we get

$$T(\bar{v}) = c_1T(\bar{v}_1) + \cdots + c_nT(\bar{v}_n) \tag{94}$$

in other words, $T(\bar{v})$ is in spans$(T(\mathcal{B}))$, as required.

Comment to the Theorem 6.15: The Theorem 6.15 applies, in partivular, when $\mathcal{B}$ is a basis for V. It could be nice if $T(\mathcal{B})$ would then be a basis for the range of T, but it is not always the case (see section 6.5 in the book.)

Composition of LT

If $T : U \to V$ and $S : V \to W$ are LT, then the composition of S with T is the mapping $S \circ T$, defined by

$$(S \circ T)(\bar{u}) = S(T(\bar{u})) \tag{95}$$

where $\bar{u}$ is in U.

Example 6.56: Let $T : \mathbb{R}^2 \to \mathcal{P}_1$ and $S : \mathcal{P}_1 \to \mathcal{P}_2$ be the LT defined by

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a+b)x \tag{96}$$

$$S(p(x)) = xp(x) \tag{97}$$

find

$$(S \circ T)\begin{bmatrix} 3 \\ -2 \end{bmatrix} \tag{98}$$

We should get $3x + x^2$.

Theorem 6.16: If $T : U \to V$ and $S : V \to W$ are LT, then $S \circ T : U \to W$ is a LT.
Proof: Let $\bar{u}$ and $\bar{v}$ be in U and let C be a scalar. Then

$$(S \circ T)(\bar{u} + \bar{v}) = S(T(\bar{u} + \bar{v})) = S(T(\bar{u}) + T(\bar{v})) = S(T(\bar{u})) + S(T(\bar{v}) = (S \circ T)(\bar{u}) + (S \circ T)(\bar{v})$$

and

$$(S \circ T)(c\bar{u}) = S(T(c\bar{u})) = S(cT(\bar{u})) = cS(T(\bar{u})) = c(S \circ T)(\bar{u}) \tag{99}$$

Inverses of LT

A LT $T : V \to W$ is invertible if there is a LT $T' : W \to V$ such that

$$T' \circ T = I_V \quad \text{and} \quad T \circ T' = I_W \tag{100}$$

In this case, T' is called an inverse for T.

Remarks:

- The domain V and codomain W of T do not have to be the same, as they do in the case of invertible matrix transformations. However, we will see in the next section that V and W must be very closed related.
- The requirement that T' be linear could have been omitted from this definition. For, as we will see in Theorem 6.24, it T' is any mapping from W to V such that $T' \circ T = I_V$ and $T \circ T' = I_W$, then T' is forced to be linear as well.
- If T' is an inverse for T, then the definition implies that T is a inverse for T'. Hence, T' is invertible to.

Example 6.58: The mapping $T : \mathbb{R}^2 \to \mathcal{P}_1$ and $T' : \mathcal{P}_1 \to \mathbb{R}^2$ are inverses. Where

$$T \begin{bmatrix} a \\ b \end{bmatrix} = a + (a+b)x \tag{101}$$

$$T'(c + dx) = \begin{bmatrix} c \\ d - c \end{bmatrix} \tag{102}$$

Solution:
Let us check that $T' \circ T = I_{\mathbb{R}^2}$

$$(T' \circ T) \begin{bmatrix} a \\ b \end{bmatrix} = T' \left(T \begin{bmatrix} a \\ b \end{bmatrix} \right) = T'(a + (a+b)x) = \begin{bmatrix} a \\ (a+b) - a \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

Next, we check that $T \circ T' = I_{\mathcal{P}_1}$

$$(T \circ T')(c + dx) = T(T'(c + dx)) = T \begin{bmatrix} c \\ d - c \end{bmatrix} = c + (c + (d - c))x = c + dx$$

Theorem 6.17: If T is an invertible LT, then its inverse is unique and denoted by T^{-1}.
Proof: The proof is the same as that of Theorem 3.6 of the book, with products of matrices replaced by compositions of LT (Exercise 31 in the book.)

Comments about the topic of the next two sections: In the next two sections, we will address the issue of determining when a given LT is invertible and finding its inverse when it exists.

The Kernel and Range of a Linear Transformation

This section extends the notion of null space and column space to the *kernel* and *range* of a LT.

Kernel: Let $T : V \to W$ be a LT. The kernel of T, denoted $\ker(T)$, is the set of all vectors in V that are mapped by T to $\bar{0}$ in W. That is

$$ker(T) = \{\bar{v} \in V : T(\bar{v}) = \bar{0}\} \tag{103}$$

Range: The range of T, denoted range(T), is the set of all vectors in W that are images of vectors in V under T. That is,

$$range(T) = \{T(\bar{v}) : \bar{v} \in V\} \quad (104)$$
$$= \{\bar{w} \in W : \bar{w} = T(\bar{v}) \text{ for some } \bar{v} \in V\} \quad (105)$$

Example 6.59: Let A be an $m \times n$ matrix and let $T = T_A$ be the corresponding matrix transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined by $T(\bar{v}) = A\bar{v}$. Then

- the range of T is the column space of A (see Chapter 3 of the book.)
- the kernel of T is the null(A)

Exercise (Example 6.60): Find the kernel and range of the differential operator $D : \mathcal{P}_3 \to \mathcal{P}_2$ defined by $D(p(x)) = p'(x)$.
Solution:

- Kernel:

$$ker(D) = \{a_0 + a_1x + a_2x^2 + a_3x^3 : D(a_0 + a_1x + a_2x^2 + a_3x^3) = 0\}$$
$$= \{a_0 + a_1x + a_2x^2 + a_3x^3 : a_1 + 2a_2x + 3a_3x^2 = 0\}$$

  dado que $a_1 + 2a_2x + 3a_3x^2 = 0$ tiene que verificarse para todo x, debe ser $a_1 = 0$, $a_2 = 0$, $a_3 = 0$, quedando a_0 indeterminado. Luego,

$$ker(D) = \{a_0 : a_0 \in \mathbb{R}\}$$

  i.e. constant polynomials (in $\mathcal{P}_3$.)

- Range(D)=$\mathcal{P}_2$

Exercise for the student in class (Example 6.61): Find the kernel and range of the integral operator $S : \mathcal{P}_1 \to \mathbb{R}$ with

$$S(p(x)) = \int_0^1 p(x)dx \quad (106)$$

Solution:

- Kernel:

$$ker(S) = \{f(x) = a + bx \in \mathcal{P}_1 : S(p(x)) = \int_0^1 f(x)dx = 0\}$$
$$= \{a + bx \in \mathcal{P}_1 : \int_0^1 (a + bx)dx = 0\}$$
$$= \{a + bx \in \mathcal{P}_1 : a + \frac{b}{2} = 0\}$$
$$\Rightarrow b = -2a$$
$$ker(D) = \{a(1 - 2x) \in \mathcal{P}_1\}$$

- Range(S)=$\mathbb{R}$ since every real number can be obtained as the image under S of some polynomial in $\mathcal{P}_1$, example, for arbitrary $a \in \mathbb{R}$, we have $\int_0^1 adx = a$.

Theorem 6.18: Let $T : V \to W$ be a LT. Then

a. The kernel of T is a subspace of V.

b. The range of T is a subspace of W.

Proof: (a) Since $T(\bar{0}) = \bar{0}$, the zero vector of V is in $\ker(T)$, so $\ker(T)$ is nonempty. Let $\bar{u}$ and $\bar{v}$ be in $\ker(T)$ and let c be a scalar. Then $T(\bar{u}) = T(\bar{v}) = \bar{0}$, so

$$T(\bar{u}+\bar{v}) = T(\bar{u}) + T(\bar{v}) = T(\bar{0}) + T(\bar{0}) = \bar{0} \quad (107)$$
$$T(c\bar{u}) = cT(\bar{u}) = c\bar{0} = \bar{0} \quad (108)$$

(b) see book, pag. 488

Definitions: Let $T : V \to W$ be a LT. Then

Rank: the rank of T is the dimension of the range of T, denoted by rank(T).

Nullity: The nullity of T is the dimension of the kernel of T and is denoted by nullity(T).

Example:

- If A is a matrix and $T = T_A$ is the matrix transformation defined by $T(\bar{v}) = A\bar{v}$, then the range and kernel of T are the column space and the null space of A, respectively: rank(T)=rank(A) and nullity(T)=nullity(A)
- The rank and nullity of the LT $D(p(x) = p'(x)$, $D : \mathcal{P}_3 \to \mathcal{P}_2$ are: Since the range(D)=$\mathcal{P}_2$ (see above) we have rank(D)=dim $\mathcal{P}_2$=3. While the kernel of D is the set of all constant polynomials: ker(D)=$\{a : a \in \mathbb{R}\} = \{a \cdot 1 : a \in \mathbb{R}\}$, hence $\{1\}$ is a basis for ker(D), son nullity(D)=dim(ker(D))=1.
- The rank and nullity of $S : \mathcal{P}_1 \to \mathbb{R}$ with $S(p(x)) = \int_0^1 p(x)dx$ are: since range(S)=$\mathbb{R}$, then rank(S)=dim(R)=1. The kernel ker(S)=$\{a(1-2x) : a \in \mathbb{R}\} = span(1-2x)$, the $1-2x$ is a basis for ker(S). Therefor, nullity(S)=dim(ker(S))=1.

The Rank Theorem (Theorem 6.19): Let $T : V \to W$ be a LT from a finite-dimensional vector space V into a vector space W. Then

$$rank(T) + nullity(T) = dim(V) \quad (109)$$

Proof: See book, pag. 490.

Examples: Using the information from the above examples we have

- $D(p(x) = p'(x)$, $D : \mathcal{P}_3 \to \mathcal{P}_2$

$$rank(D) + nullity(D) = 3 + 1 = 4 = dim(\mathcal{P}_3) \quad (110)$$

- $S : \mathcal{P}_1 \to \mathbb{R}$ with $S(p(x)) = \int_0^1 p(x)dx$

$$rank(S) + nullity(S) = 1 + 1 = 2 = dim(\mathcal{P}_1) \quad (111)$$

Exercise (Example 6.67): Find the rank and nullity of the LT $T : \mathcal{P}_2 \to \mathcal{P}_3$ define by $T(p(x)) = xp(x)$. (firs check that T is linear)
Solution:

$$\begin{aligned} ker(T) &= \{a + bx + cx^2 \in \mathcal{P}_2 : T(a + bx + cx^2) = 0\} \\ &= \{a + bx + cx^2 : ax + bx^2 + cx^3 = 0\} \\ &\quad \Rightarrow a = b = c = 0 \\ ker(T) &= \{0\} \end{aligned}$$

so we have nullity(T)=dim(ker(T))=0. The Rank Theorem implies that

$$rank(T) = dim(\mathcal{P}_2) - nullity(T) = 3 - 0 = 3 \tag{112}$$

Exercise for the student in class (Example 6.68): Let W be the vector space of all symmetric 2×2 matrices. Define a LT $T : W \to \mathcal{P}_2$ by (check that it is linear)

$$T \begin{bmatrix} a & b \\ b & c \end{bmatrix} = (a - b) + (b - c)x + (c - a)x^2 \tag{113}$$

Find the rank and nullity of T.
Solution: we start calculating the nullity because it is easier that to calculate the rank:

$$ker(T) = \left\{ \begin{bmatrix} a & b \\ b & c \end{bmatrix} : T \begin{bmatrix} a & b \\ b & c \end{bmatrix} = 0 \right\} \tag{114}$$

$$= \left\{ \begin{bmatrix} a & b \\ b & c \end{bmatrix} : (a - b) + (b - c)x + (c - a)x^2 = 0 \right\} \tag{115}$$

$$\Rightarrow (a - b) = 0,\ (b - c) = 0\,(c - a) = 0 \tag{116}$$

$$\Rightarrow a = b,\ b = c\, c = a \Rightarrow a = b = c \tag{117}$$

$$ker(T) = \left\{ \begin{bmatrix} a & a \\ a & a \end{bmatrix} \right\} = span \left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right) \tag{118}$$

then nullity(T)=dim(ker(T))=1.

Finally, from the Rank Theorem: rank(T)=dim W -nullity(T)=3-1=2.

One-to-One and Onto Linear Transformation

This section deals with the criteria for a LT to be invertible.

One-to-one transformation: A LT $T : V \to W$ is called one-to-one if T maps distinct vectors in V to distinct vectors in W another way to say is

- $T : V \to W$ is one-to-one if, for all $\bar{u}$ and $\bar{v}$ in V, $\bar{u} \neq \bar{v}$ implies that $T(\bar{u}) \neq T(\bar{v})$ or
- $T : V \to W$ is one-to-one if, for all $\bar{u}$ and $\bar{v}$ in V, $T(\bar{u}) = T(\bar{v})$ implies that $\bar{u} = \bar{v}$

Onto transformation: A LT $T : V \to W$ is called onto if range(T)=W. Another way to say it is: $T : V \to W$ is onto if, for all $\bar{w}$ in W, there is at least one $\bar{v}$ in V such that $\bar{w} = T(\bar{v})$.

Examples: Let us consider $T : \mathbb{R}^2 \to \mathbb{R}^3$ define by

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ x - y \\ 0 \end{bmatrix} \tag{119}$$

one-to-one : It is one-to-one as we show below. Let

$$T\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = T\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \tag{120}$$

then

$$\begin{bmatrix} 2x_1 \\ x_1 - y_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2x_2 \\ x_2 - y_2 \\ 0 \end{bmatrix} \tag{121}$$

so $x_1 = x_2$ and $x_1 - y_1 = x_2 - y_2$, then $y_1 = y_2$. Hence $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$, so T is one-to-one.

onto: T is not onto, since its range is not all of $\mathbb{R}^3$.

Theorem 6.20: A LT $T : V \to W$ is one-to-one if and only if ker(T)=$\{\bar{0}\}$.
Proof: See book, pag. 494.

Example 6.70: Show that the LT $T : \mathbb{R}^2 \to \mathcal{P}_1$ define by

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a + b)x \tag{122}$$

is one-to-one and onto.
Solution:
The kernel of T is $\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ since

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a + b)x = 0 \Rightarrow a = b = 0 \tag{123}$$

Consequently, T is one-to-one by the Theorem 6.20.
By the Rank Theorem, rank(T)=dim($\mathbb{R}^2$) $- \textit{nullity}(T) = 2 - 0 = 2$. Therefore, the range of T is a two-dimensional subspace of $\mathbb{R}^2$, and hence range(T)=$\mathbb{R}^2$. It follows that T is onto.

Theorem 6.21: Let dim$V$=dim$W$=n. The a LT $T : V \to W$ is one-to-one if and only if it is onto.
Proof: See book, pag. 494.

Theorem 6.22: Let $T : V \to W$ be a one-to-one LT. If $S = \{\bar{v}_1, \cdots, \bar{v}_k\}$ is LI set in V, then $T(S) = \{T(\bar{v}_1), \cdots, T(\bar{v}_k)\}$ is a LI set in W.
Proof: See book, pag. 495.

Corollary 6.23: Let dim$V$=dim$W$=n. Then a one-to-one LT: $T : V \to W$ maps a basis for V to a basis for W.
Proof: See book, pag. 495.

Example 6.71: Let $T : \mathbb{R}^2 \to \mathcal{P}_1$ be a LT define by

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a+b)x \tag{124}$$

Then, by Corollary 6.23, the standard basis $\mathcal{E} = \{\bar{e}_1, \bar{e}_2\}$ for $\mathbb{R}^2$ is mapped to a basis $T(\mathcal{E}) = \{T(\bar{e}_1), T(\bar{e}_2)\}$ of $\mathcal{P}_1$. We find that

$$T(\bar{e}_1) = T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 + (1+0)x = 1 + x \tag{125}$$

$$T(\bar{e}_2) = T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 + (0+1)x = x \tag{126}$$

$$\tag{127}$$

It follows that $\{1 + x, x\}$ is a basis for $\mathcal{P}_1$.

Theorem 6.24: A LT $T : V \to W$ is invertible if and only if it is one-to-one and onto.
Proof: See book, pag. 496.

Isomorphisms of Vector Spaces

A LT $T : V \to W$ is called an isomorphisms if it is one-to-one and onto. If V and W are two vector spaces such that there is an isomorphism from V to W, then we say that V is isomorphic to W and write $V \cong W$.

Examples: The following vector spaces are isomorphic each other:

- $\mathcal{P}_n$ and $\mathbb{R}^n$ (think about the definition of coordinates) (Example 6.72 in the book, pag. 497)
- M_{mn} and $\mathbb{R}^{mn}$ (think about the definition of coordinates) (Example 6.73 in the book, pag. 497)

The easiest way to check if two vector spaces are isomorphic is simply to check their dimensions (see the next theorem)

Theorem 6.25: Let V and W be two finite-dimensional vector spaces. Then V is isomorphic to W if and only if dim$V$=dim$W$.
Proof: See book, pag. 498.

Examples:

- $\mathcal{P}_n$ and $\mathbb{R}^n$ are no isomorphic since, dim($\mathcal{P}_n$)=n+1$\neq$dim($\mathbb{R}^n$)=n.
- The vector space of symmetric matrices M_{22}^{sym} is isomorphic with $\mathbb{R}^3$, since dim(M_{22}^{sym})=dim($\mathbb{R}^3$)=3 (see above previous examples when we calculate basis in order to determine the dimension of a basis and spaces)

The Matrix of a Linear Transformation

In this section it is shown that every LT between finite-dimensional vector spaces can be represented as a matrix transformation.

Suppose that V is an n-dimensional vector space, W is an m-dimensional vector space, and $T : V \to W$ is a LT. Let $\mathcal{B}$ and $\mathcal{C}$ be bases for V and W, respectively. Then the coordinate vector mapping $R(\bar{v}) = [\bar{v}]_{\mathcal{B}}$ defines an isomorphism $R : V \to \mathbb{R}^n$. At the same time, we have an isomorphism $S : W \to \mathbb{R}^m$ given by $S(\bar{w}) = [\bar{w}]_{\mathcal{C}}$, which allows us to associate the image $T(\bar{v})$ with the vector $[T(\bar{v})]_{\mathcal{C}}$ in $\mathbb{R}^m$ (see Fig. 2 Since R is an isomorphism, it is invertible, so

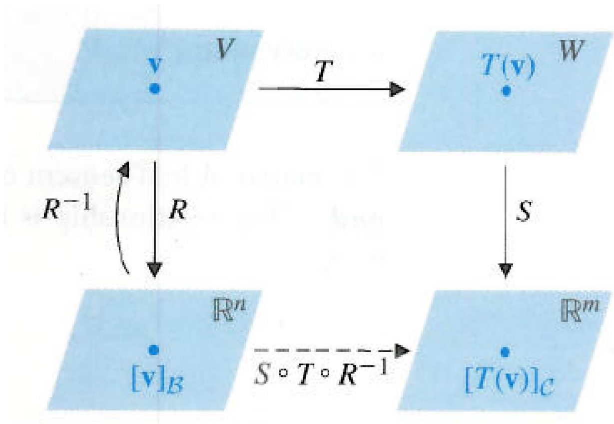

Figure 2: (from the book)

we may form the composite mapping

$$S \circ T \circ R^{-1} : \mathbb{R}^n \to \mathbb{R}^m \tag{128}$$

which maps $[\bar{v}]_{\mathcal{B}}$ to $[T(\bar{v})]_{\mathcal{C}}$. Since this mapping goes from $\mathbb{R}^n$ to $\mathbb{R}^m$, **this is a matrix transformation**.

We would like to find the $m \times n$ matrix A such that

$$A[\bar{v}]_{\mathcal{B}} = (S \circ T \circ R^{-1})[\bar{v}]_{\mathcal{B}} \tag{129}$$

$$= [T(\bar{v})]_{\mathcal{C}} \tag{130}$$

Where the columns of A are the images of the standard basis vectors for $\mathbb{R}^n$ under $S \circ T \circ R^{-1}$. But, if $\mathcal{B} = \{\bar{v}_1, \cdots, \bar{v}_n\}$ is a basis for V, then

$$R(\bar{v}_i) = [\bar{v}_i]_{\mathcal{B}} \tag{131}$$

$$= \begin{bmatrix} 0 \\ \vdots \\ 1 \vdots \\ 0 \end{bmatrix} \tag{132}$$

$$= \bar{e}_i \tag{133}$$

so

$$R^{-1}(\bar{e}_i) = \bar{v}_i \tag{134}$$

Therefore, the ith column of the matrix A we seek is given by

$$(S \circ T \circ R^{-1})(\bar{e}_i) = S(T(R^{-1}(\bar{e}_i))) \tag{135}$$

$$= S(T(\bar{v}_i)) \tag{136}$$

$$= [T(\bar{v})]_{\mathcal{C}} \tag{137}$$

which is the coordinate vector of $T(\bar{v})$ with respect to the basis $\mathcal{C}$ of W.

Theorem 6.26: Let V and W be two finite-dimensional vector spaces with bases $\mathcal{B}$ and $\mathcal{C}$, respectively, where $\mathcal{B} = \{\bar{v}_1, \cdots, \bar{v}_n\}$. If $T: V \to W$ is a LT, then the $m \times n$ matrix A defined by

$$A = [[T(\bar{v}_1)]_{\mathcal{C}}| \cdots |[T(\bar{v}_n)]_{\mathcal{C}}] \tag{138}$$

satisfies

$$A[\bar{v}]_{\mathcal{B}} = [T(\bar{v})]_{\mathcal{C}} \tag{139}$$

for every vector $\bar{v}$ in V.

Where the matrix A is called **matrix of T with respect to the bases $\mathcal{B}$ and $\mathcal{C}$**. The relationship is illustrated in diagram of fig. 3 called commutative diagram,

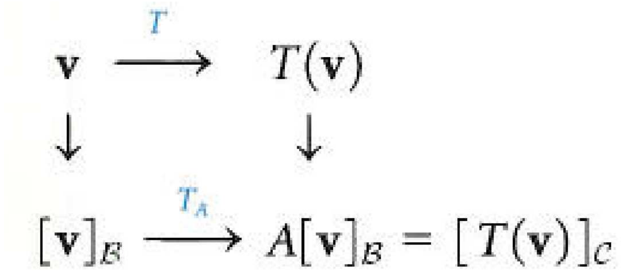

Figure 3: Commutative diagram (from the book).

Remarks:

- The matrix of a LT T with respect to bases $\mathcal{B}$ and $\mathcal{C}$ is sometimes denoted by $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$
- The matrix of a LT with respect to given basis is unique. That is, for every vector $\bar{v}$ in V, there is only one matrix A whit the property specified by Theorem 6.26–namely, $A[\bar{v}]_{\mathcal{B}} = [T(\bar{v})]_{\mathcal{C}}$
- The matrix $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ depends on the order of the vectors in the bases $\mathcal{B}$ and $\mathcal{C}$ (remember the exercises done earlier)

Exercise (Example 6.76): Let $T: \mathbb{R}^3 \to \mathbb{R}^2$ be the LT defined by

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - 2y \\ x + y - 3z \end{bmatrix} \tag{140}$$

and let $\mathcal{B} = \{\bar{e}_1, \bar{e}_2, \bar{e}_3\}$ and $\mathcal{C} = \{\bar{e}_2, \bar{e}_1\}$ be bases for $\mathbb{R}^3$ and $\mathbb{R}^2$, respectively. Find the matrix of T with respect to $\mathcal{B}$ and $\mathcal{C}$ and verify Theorem 6.26 for

$$\bar{v} = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} \tag{141}$$

Solution:
We want to built the matrix

$$A = [T]_{\mathcal{C}\leftarrow\mathcal{B}} = [[T(\bar{v}_1)]_{\mathcal{C}}|\cdots|[T(\bar{v}_n)]_{\mathcal{C}}] \tag{142}$$

For this end we need:

a. First calculate the images $T(\bar{e}_i)$ of the vectors of the basis $\mathcal{B}$

b. Then calculates the coefficient vectors of these images with respect to the basis $\mathcal{C}$ of W

c. Finally, we built the matrix $[[T(\bar{v}_1)]_{\mathcal{C}}|\cdots|[T(\bar{v}_n)]_{\mathcal{C}}]$

[a.]

$$T(\bar{e}_1) = \begin{bmatrix}1\\1\end{bmatrix} \quad T(\bar{e}_2) = \begin{bmatrix}-2\\1\end{bmatrix} \quad T(\bar{e}_3) = \begin{bmatrix}0\\-3\end{bmatrix} \tag{143}$$

[b.] Be careful with the order of the basis vectors in $\mathcal{C}$!

$$T(\bar{e}_1) = \begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix} + \begin{bmatrix}1\\0\end{bmatrix} \tag{144}$$

$$\Rightarrow [T(\bar{v}_1)]_{\mathcal{C}} = \begin{bmatrix}1\\1\end{bmatrix} \tag{145}$$

$$T(\bar{e}_2) = \begin{bmatrix}-2\\1\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix} - 2\begin{bmatrix}1\\0\end{bmatrix} \tag{146}$$

$$\Rightarrow [T(\bar{v}_2)]_{\mathcal{C}} = \begin{bmatrix}1\\-2\end{bmatrix} \tag{147}$$

$$T(\bar{e}_3) = \begin{bmatrix}0\\-3\end{bmatrix} = -3\begin{bmatrix}0\\1\end{bmatrix} + 0\begin{bmatrix}1\\0\end{bmatrix} \tag{148}$$

$$\Rightarrow [T(\bar{v}_3)]_{\mathcal{C}} = \begin{bmatrix}-3\\1\end{bmatrix} \tag{149}$$

[c.]

$$\begin{aligned} A &= [T]_{\mathcal{C}\leftarrow\mathcal{B}} && (150)\\ &= [[T(\bar{v}_1)]_{\mathcal{C}}|\cdots|[T(\bar{v}_n)]_{\mathcal{C}}] && (151)\\ &= \begin{bmatrix}1 & 1 & -3\\ 1 & -2 & 0\end{bmatrix} && (152)\end{aligned}$$

In order to verified the Theorem we have to check that the matrix product $A[\bar{v}]_{\mathcal{B}} = [\bar{v}]_{\mathcal{C}}$. Then we have to calculate

c. The action of T on $\bar{v}$ and built it coordinate vector $T(\bar{v})$ in the bases $\mathcal{C}$

d. Calculate the product of the matrix $A = [T]_{\mathcal{C}\leftarrow\mathcal{B}}$ time the coordinates of the vector $\bar{v}$ in the basis $\mathcal{B}$: $A[\bar{v}]_{\mathcal{B}}$

e. Check that (c) and (d) give the same vector

[c.]

$$
\begin{aligned}
T(\bar{v}) &= T\left(\begin{bmatrix}1\\3\\-2\end{bmatrix}\right) && (153)\\
&= \begin{bmatrix}-5\\10\end{bmatrix} && (154)\\
&= 10\begin{bmatrix}0\\1\end{bmatrix} - 5\begin{bmatrix}1\\0\end{bmatrix} && (155)\\
&\Rightarrow\ [\bar{v}]_{\mathcal{C}} = \begin{bmatrix}10\\-5\end{bmatrix} && (156)
\end{aligned}
$$

[d.]

$$
\begin{aligned}
\bar{v} &= \begin{bmatrix}1\\3\\-2\end{bmatrix} && (157)\\
&= \begin{bmatrix}1\\0\\0\end{bmatrix} + 3\begin{bmatrix}0\\1\\0\end{bmatrix} - 2\begin{bmatrix}0\\0\\1\end{bmatrix} && (158)\\
&\Rightarrow\ [\bar{v}]_{\mathcal{B}} = \begin{bmatrix}1\\3\\-2\end{bmatrix} && (159)
\end{aligned}
$$

Next,

$$
\begin{aligned}
[T]_{\mathcal{C}\leftarrow\mathcal{B}}[\bar{v}]_{\mathcal{B}} &= \begin{bmatrix}1 & 1 & -3\\1 & -2 & 0\end{bmatrix}\begin{bmatrix}1\\3\\-2\end{bmatrix} && (160)\\
&= \begin{bmatrix}10\\-5\end{bmatrix} && (161)
\end{aligned}
$$

[e.] OK!

Matrices of Composite and Inverse Linear Transformations

Theorem 6.27: Let U, V, and W be finite dimensional vector spaces with bases $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$, respectively. Let $T : U \to V$ and $S : V \to W$ be LT. Then

$$[S \circ T]_{\mathcal{D}\leftarrow\mathcal{B}} = [S]_{\mathcal{D}\leftarrow\mathcal{C}}[T]_{\mathcal{C}\leftarrow\mathcal{B}} \qquad (162)$$

Proof: See book, pag. 508.

Example 6.81: Use matrix methods to compute

$$(S \circ T)\begin{bmatrix}a\\b\end{bmatrix} \qquad (163)$$

for the LT $T : \mathbb{R}^2 \to \mathcal{P}_1$ and $S : \mathcal{P}_1 \to \mathcal{P}_2$ where

$$
\begin{aligned}
T\begin{bmatrix}a\\b\end{bmatrix} &= a + (a+b)x && (164)\\
S(p(x)) &= xp(x) && (165)
\end{aligned}
$$

Choosing the standard bases $\mathcal{E}$, $\mathcal{E}'$ and $\mathcal{E}''$ for $\mathbb{R}^2$, $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively. These canonical basis are

$$\begin{aligned}\mathcal{E} &= \{\bar{e}_1, \bar{e}_2\} && (166)\\ \mathcal{E}' &= \{1, x\} && (167)\\ \mathcal{E}'' &= \{1, x, x^2\} && (168)\end{aligned}$$

Solution:
We need to compute

$$[S \circ T]_{\mathcal{E} \leftarrow \mathcal{E}''} = [S]_{\mathcal{E}'' \leftarrow \mathcal{E}'}[T]_{\mathcal{E}' \leftarrow \mathcal{E}} \tag{169}$$

then we need to calculate

a. $[T]_{\mathcal{E}' \leftarrow \mathcal{E}}$

b. $[S]_{\mathcal{E}'' \leftarrow \mathcal{E}'}$

c. Next, using the Theorem 6.27 we calculate the matrix $(S \circ T)_{\mathcal{E}'' \leftarrow \mathcal{E}} = [S]_{\mathcal{E}'' \leftarrow \mathcal{E}'}[T]_{\mathcal{E}' \leftarrow \mathcal{E}}$, i.e. the product of the above two matrices.

d. Using the theorem 6.26 we calculate the coordinate vector of the vector which results from the action of $(S \circ T)$:

$$\left[(S \circ T)\begin{bmatrix} a \\ b \end{bmatrix}\right]_{\mathcal{E}''} = [(S \circ T)]_{\mathcal{E}'' \leftarrow \mathcal{E}}\begin{bmatrix} a \\ b \end{bmatrix} \tag{170}$$

e. With the information of the vector coordinate of the vector in the vector space $\mathcal{P}_2$ we rebuilt the vector expanded in the basis $\mathcal{E}''$.

[a.] $[T]_{\mathcal{E}' \leftarrow \mathcal{E}}$ is the matrix generated by the coordinates vector of the images of the basis vector of $\mathcal{E}$ under la action of T expressed in the basis $\mathcal{E}'$, then

$$\begin{aligned}T(\bar{e}_1) &= 1 + (1 + 0)x = 1 + x && (171)\\ T(\bar{e}_2) &= 0 + (0 + 1)x = x && (172)\end{aligned}$$

Next we expand these images in the basis $\mathcal{E}'$

$$\begin{aligned}T(\bar{e}_1) &= 1 + x = (1)1 + (1)x \Rightarrow [T(\bar{e}_1)]_{\mathcal{E}'} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} && (173)\\ T(\bar{e}_2) &= x = (0)1 + (1)x \Rightarrow [T(\bar{e}_2)]_{\mathcal{E}'} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} && (174)\end{aligned}$$

Then

$$[T]_{\mathcal{E}' \leftarrow \mathcal{E}} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \tag{175}$$

[b.] $[S]_{\mathcal{E}''\leftarrow\mathcal{E}'}$ is the matrix generated by the coordinates vector of the imagines of the basis vector of $\mathcal{E}'$ under la action of S expressed in the basis $\mathcal{E}''$, then

$$S(1) = x(1) = x \tag{176}$$
$$S(x) = x(x) = x^2 \tag{177}$$

Next we expand these images in the basis $\mathcal{E}''$

$$S(1) = x = (0)1 + (1)x + (0)x^2 \tag{178}$$
$$\Rightarrow [S(1)]_{\mathcal{E}''} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \tag{179}$$
$$S(x) = x^2 = (0)1 + (0)x + (1)x^2 \tag{180}$$
$$\Rightarrow [S(x)]_{\mathcal{E}''} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{181}$$

Then

$$[S]_{\mathcal{E}''\leftarrow\mathcal{E}'} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{182}$$

[c.]

$$(S \circ T)_{\mathcal{E}''\leftarrow\mathcal{E}} = [S]_{\mathcal{E}''\leftarrow\mathcal{E}'}[T]_{\mathcal{E}'\leftarrow\mathcal{E}} \tag{183}$$
$$= \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \tag{184}$$
$$= \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \tag{185}$$

[d.]

$$\left[(S \circ T)\begin{bmatrix} a \\ b \end{bmatrix}\right]_{\mathcal{E}''} = [(S \circ T)]_{\mathcal{E}''\leftarrow\mathcal{E}} \begin{bmatrix} a \\ b \end{bmatrix} \tag{186}$$
$$= \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \tag{187}$$
$$= \begin{bmatrix} 0 \\ a \\ a+b \end{bmatrix} \tag{188}$$

[e.] In the basis

$$\mathcal{E}'' = \{1, x, x^2\} \tag{189}$$

the vector $(S \circ T)\bar{v}$ results,

$$(S \circ T)\bar{v} = (S \circ T) \begin{bmatrix} a \\ b \end{bmatrix} \tag{190}$$

$$= (0)1 + (a)x + (a+b)x^2 \tag{191}$$

$$= ax + (a+b)x^2 \tag{192}$$

which is the same result we obtained earlier in the Example 6.56.

Theorem 6.28: Let $T : V \to W$ be a LT between n-dimensional vector spaces V and W and let $\mathcal{B}$ and $\mathcal{C}$ be bases for V and W, respectively. Then T is invertible if and only if the matrix $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ is invertible. In this case,

$$([T]_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} = [T^{-1}]_{\mathcal{B}\leftarrow\mathcal{C}} \tag{193}$$

Notice the change in the order of the basis!!
Proof: See book, pag. 509.

Example 6.82: In Example 6.70, the LT $T : \mathbb{R}^2 \to \mathcal{P}_1$ defined by

$$T \begin{bmatrix} a \\ b \end{bmatrix} = a + (a+b)x \tag{194}$$

was shown to be one-to-one and onto and hence invertible. Find T^{-1}.
Solution:
In Example 6.81, we found the matrix of T with respect to the standard bases $\mathcal{E}$ and $\mathcal{E}'$ for $\mathbb{R}^2$ and $\mathcal{P}_1$, respectively, to be

$$[T]_{\mathcal{E}'\leftarrow\mathcal{E}} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \tag{195}$$

By Theorem 6.28, it follows that the matrix of T^{-1} with respect to $\mathcal{E}'$ and $\mathcal{E}$ is

$$[T^{-1}]_{\mathcal{E}\leftarrow\mathcal{E}'} = ([T]_{\mathcal{E}'\leftarrow\mathcal{E}})^{-1} \tag{196}$$

$$= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1} \tag{197}$$

$$= \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \tag{198}$$

By theorem 6.26,

$$[T^{-1}(a+bx)]_{\mathcal{E}} = [T^{-1}]_{\mathcal{E}\leftarrow\mathcal{E}'}[a+bx]_{\mathcal{E}'} \tag{199}$$

$$= \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \tag{200}$$

$$= \begin{bmatrix} a \\ b-a \end{bmatrix} \tag{201}$$

From this coordinate vector we can reconstruct the vector using the basis $\mathcal{E}$ of $\mathbb{R}^2$

$$T^{-1}(a+bx) = a\bar{e}_1 + (b-a)\bar{e}_2 = \begin{bmatrix} a \\ b-a \end{bmatrix} \tag{202}$$

Notice that even with the last two vector looks the same they have very different mean, **clarified it by yourself!**

Change of Basis and Similarity

Theorem 6.29: Let V be a finite-dimensional vector space with bases $\mathcal{B}$ and $\mathcal{C}$ and let $T : V \to V$ be a LT. Then

$$[T]_{\mathcal{C}} = (P_{\mathcal{B}\leftarrow\mathcal{C}})^{-1}[T]_{\mathcal{B}}P_{\mathcal{B}\leftarrow\mathcal{C}} \tag{203}$$

Notice that the matrix on the most right is the change-of-basis from $\mathcal{C}$ to $\mathcal{B}$!!.
Comentario: notar que esto no es el cambio de base del vector $\bar{v}$ dentro del mismo espacio V, sino como transforma las coordenadas del operador T en dos bases diferentes dentro del mismo espacio V!!!.

Example 6.84 This example is an application of the above Theorem. The Theorem is used in order to find a basis with respect to which the matrix of LT is diagonal. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 3y \\ 2x + 2y \end{bmatrix} \tag{204}$$

If possible, find a basis $\mathcal{C}$ for $\mathbb{R}^2$ such that the matrix of T with respect to $\mathcal{C}$ is diagonal.
Solution:

a. First we write the matrix T with respect to the canonical basis $\mathcal{E}$ of $\mathbb{R}^2$

b. Search for the eigenvalues and eigenvectors

c. Then we define a new basis with the eigenvectors

[**a.**] For this end, first we calculate the action of T on each vector of the canonical basis and then we build the matrix $[T]_{\mathcal{E}}$ with the images as columns:

$$T(\bar{e}_1) = T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = (1)\bar{e}_1 + (2)\bar{e}_2 \tag{205}$$

$$\Rightarrow [T(\bar{e}_1)]_{\mathcal{E}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \tag{206}$$

$$T(\bar{e}_2) = T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 2 \end{bmatrix} = (3)\bar{e}_1 + (2)\bar{e}_2 \tag{207}$$

$$\Rightarrow [T(\bar{e}_2)]_{\mathcal{E}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \tag{208}$$

Then

$$[T]_{\mathcal{E}} = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix} \tag{209}$$

[**b.**] We search for the eigenvalues and the corresponding eigenvectors and generates the diagonal matrix D and the P in the decomposition $[T]_{\mathcal{E}} = P^{-1}DP$.
We should get

$$D = \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix} \tag{210}$$

$$P = \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix} \tag{211}$$

Then

$$P^{-1}[T]_{\mathcal{E}}P = D \tag{212}$$

[d.] Now we defined a new basis $\mathcal{C}$ form by the columns of P, i.e. the eigenvectors of $[T]_{\mathcal{E}}$. This implies that $P = P_{\mathcal{E}\leftarrow\mathcal{C}}$, i.e. the change-of-basis matrix from $\mathcal{C}$ to $\mathcal{E}$,

$$P_{\mathcal{E}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix} \tag{213}$$

then

$$(P_{\mathcal{E}\leftarrow\mathcal{C}})^{-1}[T]_{\mathcal{E}}P_{\mathcal{E}\leftarrow\mathcal{C}} = D \tag{214}$$

$$[T]_{\mathcal{C}} = D \tag{215}$$

so, the matrix T with respect to the basis

$$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \end{bmatrix} \right\} \tag{216}$$

is diagonal!!

Diagonalizable Transformation: Let V be a finite-dimensional vector space and let $T : V \to V$ be a LT. Then T is called diagonalizable if there is a basis $\mathcal{C}$ for V such that the matrix $[T]_{\mathcal{C}}$ is a diagonal matrix.

Remark: It is not hard to show that if $\mathcal{B}$ is any basis for V, then T is diagonalizable if and only if the matrix $[T]_{\mathcal{B}}$ is diagonalizable.

Theorem 4.17: Fundamental Theorem (FT) of Invertible Matrices. Version 4 of 5 Let A be an $n \times n$ matrix and let $T : V \to W$ be a LT whose matrix $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ with respect to bases $\mathcal{B}$ and $\mathcal{C}$ of V and W, respectively, is A. The following statements are equivalent:

From Version 1

a. A is invertible.

b. $A\bar{x} = \bar{b}$ has a unique solution for every $\bar{b}$ in $\mathbb{R}^n$.

c. $A\bar{x} = 0$ has only the trivial solution.

d. The reduced row echelon form of A is I_n.

e. A is a product of elementary matrices.

From Version 2

f. rank(A)=n

g. nullity(A)=0

h. The column vectors of A are LI

i. The column vectors of A span $\mathbb{R}^n$

j. The column vectors of A form a basis for $\mathbb{R}^n$

k. The row vectors of A are LI

l. The row vectors of A span $\mathbb{R}^n$

m. The row vectors of A form a basis for $\mathbb{R}^n$

From Version 3

n. $det A \neq 0$

o. 0 is not an eigenvalue of A

New statements

p. T is invertible

q. T is one-to-one

r. T is onto

s. ker(T)=$\{\bar{0}\}$

t. range(T)=W

Proof: The equivalence (q)⇔(s) is Theorem 6.20, and (r)⇔(t) is the definition of onto. Since A is $n \times n$, we must have dim$V$=dim$W$=n. From Theorem 6.21 and 6.24, we get (p)⇔(q)⇔(r). Finally, we connect the last five statements to the others by Theorem 6.28, which implies that (a)⇔(p).