UNIVERSIDAD DEL PACÍFICO
Departamento Académico de Economía
Matemáticas III (130233) - Sección A
Segundo Semestre 2017
Profesor Diego Winkelried

1 | Vector spaces

1.1 Vectors

- **Definition.** A vector is an array of values that represents, in a particular space, a magnitude or distance, and a direction.

  In general, a vector is characterized by its initial and terminal points. However, we shall focus on vectors whose initial point is the origin. Thus, the terminal point contains all required information about magnitude and direction. From this viewpoint, a vector is simply a coordinate in $\mathbb{R}^n$, the n-dimensional space.

  A vector in $\mathbb{R}^n$ has dimension n and n components or entries. For instance, $\boldsymbol{a}' = (1,2)$ belongs to $\mathbb{R}^2$ whereas $\boldsymbol{a}' = (1,2,3)$ is in $\mathbb{R}^3$.

- **Matrix representation.** Mechanically, a vector is a one-row matrix (*row vector*) or a one-column matrix (more usual, *column vector*). In general, in $\mathbb{R}^n$

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}_{n\times 1} .$$

  Alternative notation: $\boldsymbol{a}' = (a_1, a_2, \ldots, a_n)$ or $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)'$.

- **Basic operations.**

  (1) **Equality.** Two vectors in $\mathbb{R}^n$ are equal if and only if both represent the same terminal point. This happens if equality occurs on an element by element basis. That is, $\boldsymbol{a} = \boldsymbol{b}$ if and only if $a_i = b_i$ for $i = 1, 2, \ldots, n$.

  (2) **Addition or sum.** The sum of two vectors is a vector whose entries are given by the sum of the corresponding entries of the original vectors.
  Let $\boldsymbol{a} \in \mathbb{R}^n$ and $\boldsymbol{b} \in \mathbb{R}^n$. The i-th entry of vector $\boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b} \in \mathbb{R}^n$ is $c_i = a_i + b_i$ for $i = 1, 2, \ldots, n$.

  (3) **Multiplication by a scalar.** When a vector is multiplied by a scalar, the result is a vector whose entries equal the entries of the original vector times the scalar.
  Hence, $\lambda\boldsymbol{a} = (\lambda a_1, \lambda a_2, \ldots, \lambda a_n)'$, where $\lambda \in \mathbb{R}$.
  If this operation is combined with addition, we have that $\boldsymbol{c} = \boldsymbol{a} \pm \boldsymbol{b} \in \mathbb{R}^n$ implies $c_i = a_i \pm b_i$ for $i = 1, 2, \ldots, n$.

  (4) **Dot product (or scalar product or (Euclidean) inner product).** The dot product of two vectors is a scalar that equals the sum of the products of the corresponding entries. Obviously, both vectors should have the same dimension, otherwise the dot product will not be defined. In other words, we require matrix multiplication conformability.
  Consider two vectors $\boldsymbol{a} \in \mathbb{R}^n$ and $\boldsymbol{b} \in \mathbb{R}^n$. Then,

$$\boldsymbol{a}'\boldsymbol{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i \in \mathbb{R} .$$

- **Null vector.** It represents the origin, and is a vector full of zeroes: $\mathbf{0} = (0, 0, \ldots, 0)'$.

- **Properties of the basic arithmetic operations.** For $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^n$, $\boldsymbol{c} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ (scalar),

(1) Associative	$\boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c}) = (\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c}$.
(2) Commutative additive	$\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}$.
(3) Commutative multiplicative	$\boldsymbol{a}'\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{a}$.
(4) Distributive additive	$\lambda(\boldsymbol{a} + \boldsymbol{b}) = \lambda\boldsymbol{a} + \lambda\boldsymbol{b}$.
(5) Distributive multiplicative	$\boldsymbol{c}'(\boldsymbol{a} + \boldsymbol{b}) = \boldsymbol{c}'\boldsymbol{a} + \boldsymbol{c}'\boldsymbol{b}$.
(6) Distributive multiplicative	$(\lambda\boldsymbol{a})'\boldsymbol{b} = \boldsymbol{a}'(\lambda\boldsymbol{b})$.
(7) Neutral additive element	$\boldsymbol{a} + \boldsymbol{0} = \boldsymbol{a}$.
(8) Neutral multiplicative element	$\boldsymbol{a}'\boldsymbol{0} = 0$.

- **(Euclidean) Norm or modulus.** A scalar that measures the length of the vector. It is defined as

$$\|\boldsymbol{a}\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2} = \sqrt{\sum_{i=1}^{n} a_i^2}\,.$$

- **Properties of the norm.** For $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ (scalar),
  (1) $\|\boldsymbol{a}\|^2 = \boldsymbol{a}'\boldsymbol{a}$.
  (2) In general, $\|\boldsymbol{a}\| \geq 0$. $\|\boldsymbol{a}\| = 0$ if and only if $\boldsymbol{a} = \boldsymbol{0}$.
  (3) $\|\lambda\boldsymbol{a}\| = |\lambda|\,\|\boldsymbol{a}\|$.
  (4) $\|-\boldsymbol{a}\| = \|\boldsymbol{a}\|$.
  (5) If $\|\boldsymbol{a} \pm \boldsymbol{b}\| = 0$, then necessarily $\boldsymbol{a} = \mp\boldsymbol{b}$.
  (6) $\|\boldsymbol{a} \pm \boldsymbol{b}\|^2 = \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 \pm 2\boldsymbol{a}'\boldsymbol{b}$.

- **All-ones vector (sum vector).** A vector whose entries are all equal to unity: $\boldsymbol{s} = (\,1, 1, \ldots, 1\,)'$. The dot product of any vector with $\boldsymbol{s}$ gives

$$\boldsymbol{a}'\boldsymbol{s} = \boldsymbol{s}'\boldsymbol{a} = a_1 + a_2 + \cdots + a_n = \sum_{i=1}^{n} a_i\,,$$

  hence the colloquial name of "sum vector". The norm of $\boldsymbol{s}$ is $\|\boldsymbol{s}\| = \sqrt{\boldsymbol{s}'\boldsymbol{s}} = \sqrt{n}$.

- **Unit vector.** Any vector $\boldsymbol{a}$ with unit length, $\|\boldsymbol{a}\| = 1$. Any vector $\boldsymbol{b}$ can be *normalized* as $\boldsymbol{a} = \boldsymbol{b}/\|\boldsymbol{b}\|$ to render a unit vector.

- **Canonical (unit) vector or selection vector.** The vector $\boldsymbol{e}_i \in \mathbb{R}^n$ is such that its i-th entry equals one and the remaining $n-1$ entries are equal to zero. $\boldsymbol{e}_i$ is a unit vector ($\|\boldsymbol{e}_i\| = 1$) pointing in the direction of the i-th axis of a Cartesian coordinate system.

  Note that $\boldsymbol{e}_i$ is the i-th of the $n \times n$ identity matrix. Henceforth, in $\mathbb{R}^n$, there exist n (different) canonical vectors: $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$.

  The dot product of any vector with $\boldsymbol{e}_i$ gives

$$\boldsymbol{a}'\boldsymbol{e}_i = \boldsymbol{e}_i{}'\boldsymbol{a} = a_1 \cdot 0 + a_2 \cdot 0 + \cdots + a_i \cdot 1 + \cdots + a_n \cdot 0 = a_i\,,$$

  hence the colloquial name of "selection vector" (this dot product *selects* the i-the entry of $\boldsymbol{a}$).

- **Distance between two vectors.** The norm of the difference between $\boldsymbol{a} \in \mathbb{R}^n$ and $\boldsymbol{b} \in \mathbb{R}^n$:

$$D = \|\boldsymbol{a} - \boldsymbol{b}\| = \|\boldsymbol{b} - \boldsymbol{a}\| = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2}\,.$$

- **Triangle inequality.** For $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^n$ and $\boldsymbol{c} \in \mathbb{R}^n$,

$$\|\boldsymbol{a} - \boldsymbol{b}\| + \|\boldsymbol{b} - \boldsymbol{c}\| \geq \|\boldsymbol{a} - \boldsymbol{c}\| \,.$$

- **Angle between two vectors.** $\cos(\alpha) = \dfrac{\boldsymbol{a}'\boldsymbol{b}}{\|\boldsymbol{a}\|\,\|\boldsymbol{b}\|}$ o $\alpha = \arccos\left(\dfrac{\boldsymbol{a}'\boldsymbol{b}}{\|\boldsymbol{a}\|\,\|\boldsymbol{b}\|}\right)$.

- **Notes.** It follows that $\boldsymbol{a}'\boldsymbol{b} = \|\boldsymbol{a}\|\,\|\boldsymbol{b}\|\cos(\alpha)$.
  (1) The sign of $\boldsymbol{a}'\boldsymbol{b}$ is that of $\cos(\alpha)$. Thus, $\boldsymbol{a}'\boldsymbol{b} > 0$ if α is acute, and $\boldsymbol{a}'\boldsymbol{b} < 0$ if α is obtuse.
  (2) Collinearity 1: If $\alpha = 0°$, then $\cos(\alpha) = 1$. Therefore, $\boldsymbol{a}'\boldsymbol{b} = \|\boldsymbol{a}\|\,\|\boldsymbol{b}\|$.
  (3) Collinearity 2: If $\alpha = 180°$, then $\cos(\alpha) = -1$. Therefore, $\boldsymbol{a}'\boldsymbol{b} = -\|\boldsymbol{a}\|\,\|\boldsymbol{b}\|$.
  (4) Perpendicularity or *orthogonality*: If $\alpha = 90°$, then $\cos(\alpha) = 0$. Therefore, $\boldsymbol{a}'\boldsymbol{b} = 0$. With this, $\|\boldsymbol{a} \pm \boldsymbol{b}\|^2 = \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2$ for orthogonal vectors.

- **Cauchy-Schwarz inequality.** Since $-1 \leq \cos(\alpha) \leq 1$, it follows that $-\|\boldsymbol{a}\|\,\|\boldsymbol{b}\| \leq \boldsymbol{a}'\boldsymbol{b} \leq \|\boldsymbol{a}\|\,\|\boldsymbol{b}\|$. Upon squaring,

$$(\boldsymbol{a}'\boldsymbol{b})^2 \leq \|\boldsymbol{a}\|^2\|\boldsymbol{b}\|^2 \,.$$

1.2 Vector space

- **Definition.** A set $\mathcal{V}$ is said to have the structure of a vector space (on the real numbers field) if the following holds:

Closure under addition	:	if $\boldsymbol{v} \in \mathcal{V}$ and $\bar{\boldsymbol{v}} \in \mathcal{V}$, then $\boldsymbol{v} + \bar{\boldsymbol{v}} \in \mathcal{V}$.
Closure under scalar multiplication	:	if $\boldsymbol{v} \in \mathcal{V}$ and $\lambda \in \mathbb{R}$, then $\lambda\boldsymbol{v} \in \mathcal{V}$.

  Both properties may be encompassed in a single, distributive one: if $\boldsymbol{v} \in \mathcal{V}$, $\bar{\boldsymbol{v}} \in \mathcal{V}$, $\lambda_1 \in \mathbb{R}$ and $\lambda_2 \in \mathbb{R}$ then $\lambda_1\boldsymbol{v} + \lambda_2\bar{\boldsymbol{v}} \in \mathcal{V}$.

- **Examples.**
  (1) $\mathbb{R}^n$ is a vector space: if $\boldsymbol{v}_1 \in \mathbb{R}^n$ and $\boldsymbol{v}_2 \in \mathbb{R}^n$, then $\lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2 \in \mathbb{R}^n$ for any two real numbers λ_1 y λ_2.

  (2) $\mathcal{V}(a, b) = \{\, (x_1, x_2) \in \mathbb{R}^2 \mid ax_1 + bx_2 = 0 \,\}$ is a vector space.
  Note that if $\boldsymbol{v} \in \mathcal{V}$, then $av_1 + bv_2 = 0$. Similarly, if $\bar{\boldsymbol{v}} \in \mathcal{V}$, then $a\bar{v}_1 + b\bar{v}_2 = 0$. Upon summing both equalities we find that $a(v_1 + \bar{v}_1) + b(v_2 + \bar{v}_2) = 0$, so we conclude that $\boldsymbol{v} + \bar{\boldsymbol{v}} \in \mathcal{V}$: $\mathcal{V}$ is closed under addition. On the other hand, multiplying the first equality by λ gives $a\lambda v_1 + b\lambda v_2 = 0$, leading to $\lambda\boldsymbol{v} \in \mathcal{V}$: $\mathcal{V}$ is closed under scalar multiplication.

  (3) A generalization: Given $\boldsymbol{a} \in \mathbb{R}^n$, $\mathcal{V}(\boldsymbol{a}) = \{\, \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{a}'\boldsymbol{x} = 0 \,\}$ is a vector space.
  If $\boldsymbol{v} \in \mathcal{V}$ then $\boldsymbol{a}'\boldsymbol{v} = 0$ and if $\bar{\boldsymbol{v}} \in \mathcal{V}$ then $\boldsymbol{a}'\bar{\boldsymbol{v}} = 0$. Upon summing both equalities, $\boldsymbol{a}'(\boldsymbol{v} + \bar{\boldsymbol{v}}) = 0$ and the set is closed under addition. Moreover, after multiplying the first equality by λ, we obtain $\boldsymbol{a}'(\lambda\boldsymbol{v}) = 0$ and the set is also closed under scalar multiplication.

  (4) As opposed to the previous example, given $\boldsymbol{a} \in \mathbb{R}^n$ and $\delta \neq 0$, $\mathcal{V}(\boldsymbol{a}, \delta) = \{\, \boldsymbol{x} \in \mathbb{R}^n; \mid \boldsymbol{a}'\boldsymbol{x} = \delta \,\}$ *is not a vector space*. If $\boldsymbol{v} \in \mathcal{V}$ then $\boldsymbol{a}'\boldsymbol{v} = \delta$. After multiplying by $\lambda \neq 0$, $\boldsymbol{a}'(\lambda\boldsymbol{v}) = \lambda\delta \neq 0$, so closure under scalar multiplication is not satisfied. The same is true for closure under addition.

  (5) Let $\mathcal{V}$ be the set of all coordinates (x, y) located on the first quadrant of $\mathbb{R}^2$. For two points from this set, we verify that $(x_1 + x_2, y_1 + y_2)$ is also located on the first quadrant of $\mathbb{R}^2$. However, $(\lambda x, \lambda y)$ will lie on the first quadrant of $\mathbb{R}^2$ only when $\lambda > 0$, and not in a general case $\lambda \in \mathbb{R}$. Therefore, $\mathcal{V}$ *is not a vector space*.

  (6) A similar argument leads us to conclude that, given $\boldsymbol{a} \in \mathbb{R}^n$, $\mathcal{V}(\boldsymbol{a}) = \{\, \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{a}'\boldsymbol{x} > 0 \,\}$ does not have the structure of a vector space either.

- **Vector subspace.** Often, a vector space $\mathcal{W}$ is contained within another vector space $\mathcal{V}$, and that $\mathcal{W}$ is closed under addition and multiplication just as $\mathcal{V}$ is. In this case, $\mathcal{W}$ is said to be a (vector) subspace of $\mathcal{V}$.

- **Examples.**
  (1) $\mathbb{R}^m$ is a vector subspace of $\mathbb{R}^n$ for $m < n$.

  (2) Any line that passes through the origin is a vector subspace of $\mathbb{R}^2$.

  (3) Any plane through the origin is a vector subspace of $\mathbb{R}^3$.

  (4) $\mathcal{W}(\boldsymbol{x}) = \{\, \boldsymbol{x} \in \mathbb{R}^n; \boldsymbol{a} \in \mathbb{R}^n, \boldsymbol{b} \in \mathbb{R}^n \mid \boldsymbol{a}'\boldsymbol{x} = 0 \text{ y } \boldsymbol{b}'\boldsymbol{x} = 0 \,\}$.
  We showed before that the set of vectors orthogonal to $\boldsymbol{a}$, that is $\mathcal{V}(\boldsymbol{x})$ such that $\boldsymbol{a}'\boldsymbol{x} = 0$, constituted a vector space. $\mathcal{W}(\boldsymbol{x})$ contains vectors that are orthogonal to $\boldsymbol{a}$ and that are *also* orthogonal to $\boldsymbol{b}$. Clearly, $\mathcal{W} \subset \mathcal{V}$, besides satisfying the closure requirements. $\mathcal{W}$ is a vector subspace of $\mathcal{V}$.

1.3 Linear combination

- **Definition.** Consider the set of m vectors in $\mathbb{R}^n$, $\boldsymbol{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$. Vector $\boldsymbol{a}$ is a linear combination of the vectors in $\boldsymbol{V}$ if it can be expressed as a *weighted sum* of these vectors,

$$\boldsymbol{a} = \lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2 + \cdots + \lambda_m \boldsymbol{v}_m = \sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i \,,$$

  where the scalars (weights) $\lambda_1, \lambda_2, \ldots, \lambda_m$ must exist.

- **Matrix approach.** A convenient way to determine whether a vector $\boldsymbol{a} \in \mathbb{R}^n$ can be expressed as a linear combination of a set of vectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ is to find the weights $\lambda_1, \lambda_2, \ldots, \lambda_m$ as the solution of a system of linear equations. Thus, if we define an $n \times m$ matrix whose i-th column is vector $\boldsymbol{v}_i$, $\mathbf{V} = [\, \boldsymbol{v}_1 \; \boldsymbol{v}_2 \; \cdots \; \boldsymbol{v}_m \,]$ and a vector $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_m)'$, such system is

$$\mathbf{V}\boldsymbol{\lambda} = \boldsymbol{a} \,.$$

  This system can be solved using the **Gauss-Jordan elimination method.**

- **Some cases:**
  (1) Any vector in $\mathbb{R}^n$ can be expressed as a linear combination of the n canonical unit vectors.

  Proof. Note that, by construction,

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + a_i \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + a_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \end{bmatrix} .$$

  Compactly,

$$\boldsymbol{a} = a_1 \boldsymbol{e}_1 + a_2 \boldsymbol{e}_2 + \cdots + a_i \boldsymbol{e}_i + \cdots + a_n \boldsymbol{e}_n = \sum_{i=1}^{n} a_i \boldsymbol{e}_i \,.$$

  Thus, $\boldsymbol{a}$ is a linear combination of $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m\}$, with weights given by $\lambda_1 = a_1, \lambda_2 = a_2, \ldots, \lambda_n = a_n$.

  (2) The null vector can always be expressed as a linear combination of vectors.

  Proof. Trivially, $\mathbf{0} = \sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i$, at least for $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0$.

(3) If $\boldsymbol{a}$ is a linear combination of a set of vectors $\boldsymbol{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$, it is also a linear combination of any set that contains the vectors in $\boldsymbol{V}$.

Proof. We begin by noting that there exist weights $\lambda_1, \lambda_2, \ldots, \lambda_m$ such that

$$\boldsymbol{a} = \sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i \,.$$

Consider now an augmented set of vectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m, \boldsymbol{v}_{m+1}, \boldsymbol{v}_{m+2}, \ldots, \boldsymbol{v}_p\}$. Hence,

$$\boldsymbol{a} = \sum_{i=1}^{p} \tilde{\lambda}_i \boldsymbol{v}_i = \sum_{i=1}^{m} \tilde{\lambda}_i \boldsymbol{v}_i + \sum_{i=m+1}^{p} \tilde{\lambda}_i \boldsymbol{v}_i \,.$$

This is true, at the very least, when $\tilde{\lambda}_i = \lambda_i$ for $i \leq m$ and $\tilde{\lambda}_i = 0$ for $i > m$.

1.4 Linear independence

- **Definition.** The set of vectors $\boldsymbol{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ is *linearly independent* if *none* of these vectors can be obtained as a linear combination of the remaining vectors. On the contrary, the set $\boldsymbol{V}$ is linearly dependent if *at least one* of the vectors can be expressed as a linear combination of the rest.

- **Linear contrast.** The vectors in $\boldsymbol{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ are linearly independent if the equation (known as a *linear contrast*)

$$\sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i = \mathbf{0}$$

is only satisfied by the *trivial solution* $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0$. If at least one $\lambda_i \neq 0$, then the set $\boldsymbol{V}$ is linearly dependent.

- **Matrix criterion.** The linear contrast forms a linear system of n equations (the dimension of the vectors in $\boldsymbol{V}$) and m unknowns (the scalars $\lambda_1, \lambda_2, \ldots, \lambda_m$). Thus, if we define an $n \times m$ matrix whose i-th column is vector $\boldsymbol{v}_i$, $\mathbf{V} = [\, \boldsymbol{v}_1 \; \boldsymbol{v}_2 \; \cdots \; \boldsymbol{v}_m \,]$ and a vector $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_m)'$, the aforementioned criterion amounts to

$$\mathbf{V}\boldsymbol{\lambda} = \mathbf{0} \,.$$

The vectors arrayed as the columns of $\mathbf{V}$ are, therefore, linearly independent only when this homogeneous system of linear equations is uniquely solved by $\boldsymbol{\lambda} = \mathbf{0}$. This is the case where $\text{rank}(\mathbf{V}) = m$.

- **Some special cases:**

(1) A set of m vectors in $\mathbb{R}^n$ is, necessarily, linearly dependent if $m > n$.

Proof. Consider the matrix system $\mathbf{V}\boldsymbol{\lambda} = \mathbf{0}$. In this case, we have more unknowns than equations $m > n$ (matrix $\mathbf{V}$ has more columns than rows, so $\text{rank}(\mathbf{V}) < m$ necessarily). Under such circumstances the system contains "'free variables" and has *infinitely many* solutions, so we will always be able to find a vector $\boldsymbol{\lambda} \neq \mathbf{0}$ that solve it.

(2) Any set of vectors that contains the null vector is linearly dependent.

Proof. Consider the set $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{m-1}, \mathbf{0}\}$. Thus,

$$\sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i = \sum_{i=1}^{m-1} \lambda_i \boldsymbol{v}_i + \lambda_m \mathbf{0}$$

will be equal to the null vector if $\lambda_1 = \lambda_2 = \cdots = \lambda_{m-1} = 0$ *regardless* of the value of λ_m, since $\lambda_m \mathbf{0} = \mathbf{0}$ in any case.

(3) If a set of vectors is linearly independent, any of their subsets will also be linearly independent.

Proof. Let us split the set $\boldsymbol{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ into two groups $\boldsymbol{V}_1$ y $\boldsymbol{V}_2$. Since $\boldsymbol{V}$ is linearly independent, we have that

$$\sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i = \sum_{i \in \boldsymbol{V}_1} \lambda_i \boldsymbol{v}_i + \sum_{i \in \boldsymbol{V}_2} \lambda_i \boldsymbol{v}_i = \mathbf{0}$$

only for $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0$. Suppose now that $\boldsymbol{V}_1$ es linearly *dependent*. This implies that we can find weights $\pi_i \neq 0$ (at least in one instance) such that

$$\sum_{i \in \boldsymbol{V}_1} \pi_i \boldsymbol{v}_i = \mathbf{0}, \quad \text{but this implies that} \quad \sum_{i \in \boldsymbol{V}_1} \pi_i \boldsymbol{v}_i + \sum_{i \in \boldsymbol{V}_2} 0 \cdot \boldsymbol{v}_i = \mathbf{0}$$

and that $\boldsymbol{V}$ is, as a whole, linearly dependent. A contradiction: $\boldsymbol{V}_1$ must be linearly *independent*.

(4) If a set of vectors is linearly dependent, any superset that contains it will also be linearly dependent.

Proof. Let us split $\boldsymbol{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ into two groups $\boldsymbol{V}_1$ y $\boldsymbol{V}_2$. Now, recall that $\boldsymbol{V}_1$ is linearly dependent. This means that there exists at least one $\lambda_i \neq 0$ such that

$$\sum_{i \in \boldsymbol{V}_1} \lambda_i \boldsymbol{v}_i = \mathbf{0}.$$

For the superset $\boldsymbol{V}$, we can write

$$\sum_{i \in \boldsymbol{V}_1} \lambda_i \boldsymbol{v}_i + \sum_{i \in \boldsymbol{V}_2} 0 \cdot \boldsymbol{v}_i = \mathbf{0}$$

which will be satisfied at least for one $\lambda_i \neq 0$. Therefore, $\boldsymbol{V}$ is also linearly dependent.

(5) Every set $\boldsymbol{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ of mutually orthogonal vectors (i.e., $\boldsymbol{v}_i{}'\boldsymbol{v}_j = 0$ for all $i \neq j$) is linearly independent.

Proof. Consider the linear contrast

$$\lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2 + \lambda_3 \boldsymbol{v}_3 + \cdots + \lambda_i \boldsymbol{v}_i + \cdots + \lambda_m \boldsymbol{v}_m = \mathbf{0}.$$

Upon premultiplying by $\boldsymbol{v}_i{}'$, we obtain

$$\lambda_1(\boldsymbol{v}_i{}'\boldsymbol{v}_1) + \lambda_2(\boldsymbol{v}_i{}'\boldsymbol{v}_2) + \lambda_3(\boldsymbol{v}_i{}'\boldsymbol{v}_3) + \cdots + \lambda_i(\boldsymbol{v}_i{}'\boldsymbol{v}_i) + \cdots + \lambda_m(\boldsymbol{v}_i{}'\boldsymbol{v}_m) = 0,$$

which boils down to $\lambda_i \|\boldsymbol{v}_i\|^2 = 0$. Since $\|\boldsymbol{v}_i\|^2 > 0$, we must conclude that $\lambda_i = 0$ for all i. It follows that the set is linearly independent.

1.5 Spanning vector spaces

- **Spanning set (or, simply, span).** The vectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ that belong to the vector space $\mathcal{V}$ form a spanning set if any vector in this space can be expressed as a linear combination of the vectors in this set. Formally, if $\boldsymbol{a} \in \mathcal{V}$ and $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ span (or *generate*) $\mathcal{V}$, then we can write

$$\boldsymbol{a} = \sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i.$$

As a matter of convention, it is said that $\mathcal{V}$ is *spanned* (or *generated*) by $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$. Alternatively, that the *span* of $\mathcal{V}$ is $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$.

- **Basis.** A set of vectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ that belong to $\mathcal{V}$ form a basis of $\mathcal{V}$ if:

(i) $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ span $\mathcal{V}$, and

(ii) $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ is linearly independent.

Intuitively, a basis can be thought of as an "efficient" (or the *smallest*) spanning set.

- **Dimension of the vector space.** $\dim(\mathcal{V}) = m$, is the number of vectors that form the basis.

- **Canonical basis in $\mathbb{R}^n$.** The set $V = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$. This set spans $\boldsymbol{a} \in \mathbb{R}^n$. Moreover, these are clearly linearly independent vectors.

- **Orthogonal basis in $\mathbb{R}^n$.** A basis of n mutually orthogonal vectors. We showed that such a set is linearly dependent and the property (4) below indicates that it spans $\mathbb{R}^n$.

- **Orthonormal basis in $\mathbb{R}^n$.** An orthogonal basis such that $\|\boldsymbol{v}_i\| = 1$ for all i.

- **Matrix criterion.** A convenient approach to compute a basis of $\text{span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m)$ is to find the **reduced row echelon form** of the matrix $\mathbf{V} = [\, \boldsymbol{v}_1 \; \boldsymbol{v}_2 \; \cdots \; \boldsymbol{v}_m \,]$ through the Gauss-Jordan elimination method. Then, the basis is formed by the subset of $d \leq m$ vectors of V corresponding to the columns with pivots of the reduced form of $\mathbf{V}$. In fact, $\dim(\text{span}(\mathbf{V})) = \text{rank}(\mathbf{V}) = d$.

- **Properties.**

(1) Any vector in $\mathcal{V}$ can be expressed as a *unique* linear combination of its basis.

Proof. Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ be the basis of $\mathcal{V}$. Let also $\boldsymbol{a} \in \mathcal{V}$ such that

$$\boldsymbol{a} = \sum_{i=1}^{m} \pi_i \boldsymbol{v}_i \,.$$

Suppose now that $\boldsymbol{a}$ can be further expressed as an alternative linear combination of the basis

$$\boldsymbol{a} = \sum_{i=1}^{m} \delta_i \boldsymbol{v}_i \,.$$

After subtracting both equalities, we get

$$\boldsymbol{a} - \boldsymbol{a} = \sum_{i=1}^{m} (\delta_i - \pi_i) \boldsymbol{v}_i \qquad \rightarrow \qquad \sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i = \mathbf{0}\,,$$

where $\lambda_i = \delta_i - \pi_i$. However, since the basis is linearly independent, it must be the case that $\lambda_i = 0$ for all i. We conclude that $\delta_i = \pi_i$. The linear combination is indeed unique.

(2) Let $V = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ be the basis of $\mathcal{V}$. Consider a vector $\boldsymbol{a} \in \mathcal{V}$ such that

$$\boldsymbol{a} = \sum_{i=1}^{m} \beta_i \boldsymbol{v}_i \quad \text{with} \quad \beta_r \neq 0 \quad \text{for } 1 \leq r \leq m\,.$$

Then, the vectors $V^* = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{r-1}, \boldsymbol{a}, \boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_m\}$ also form a basis of $\mathcal{V}$.

Proof. Since $\beta_r \neq 0$ we can write

$$\boldsymbol{a} = \sum_{i=1}^{m} \beta_i \boldsymbol{v}_i \quad \rightarrow \quad v_r = \frac{1}{\beta_r}\left(\boldsymbol{a} - \sum_{i \neq r} \beta_i \boldsymbol{v}_i\right).$$

On the other hand, any vector $\boldsymbol{c} \in \mathcal{V}$ can be expressed as a linear combination of the basis V. Putting together this fact with the expression for $\boldsymbol{v}_r$,

$$\boldsymbol{c} = \sum_{i=1}^{m} \pi_i \boldsymbol{v}_i = \sum_{i \neq r} \pi_i \boldsymbol{v}_i + \pi_r \boldsymbol{v}_r = \sum_{i \neq r} \pi_i \boldsymbol{v}_i + \frac{\pi_r}{\beta_r}\left(\boldsymbol{a} - \sum_{i \neq r} \beta_i \boldsymbol{v}_i\right) = \sum_{i \neq r}\left(\pi_i - \frac{\pi_r}{\beta_r}\beta_i\right)\boldsymbol{v}_i + \left(\frac{\pi_r}{\beta_r}\right)\boldsymbol{a}\,,$$

which shows that $\boldsymbol{c}$ can be written as a linear combination of $\boldsymbol{V}^*$. Thus, $\boldsymbol{V}^*$ spans $\mathcal{V}$.
We now need to show that $\boldsymbol{V}^*$ is linearly independent. Since $\boldsymbol{V}$ is linearly independent,

$$\sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i = \sum_{i \neq r} \lambda_i \boldsymbol{v}_i + \lambda_r \boldsymbol{v}_r = \mathbf{0} \tag{*}$$

only for $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0$. In the case of $\boldsymbol{V}^*$ we must verify that

$$\sum_{i \neq r} \tilde{\lambda}_i \boldsymbol{v}_i + \tilde{\lambda}_r \boldsymbol{a} = \mathbf{0}$$

only for $\tilde{\lambda}_1 = \tilde{\lambda}_2 = \cdots = \tilde{\lambda}_m = 0$. Upon replacing $\boldsymbol{a}$ by its definition,

$$\sum_{i \neq r} \tilde{\lambda}_i \boldsymbol{v}_i + \tilde{\lambda}_r \sum_{i=1}^{m} \beta_i \boldsymbol{v}_i = \sum_{i \neq r} (\tilde{\lambda}_i + \tilde{\lambda}_r \beta_i) \boldsymbol{v}_i + \tilde{\lambda}_r \beta_r \boldsymbol{v}_r = \mathbf{0}\,. \tag{**}$$

Equate (*) to (**) in order to conclude that $\tilde{\lambda}_r = \lambda_r / \beta_r = 0$ and that $\tilde{\lambda}_i = \lambda_i - \tilde{\lambda}_r \beta_i = 0$. $\boldsymbol{V}^*$ is linearly independent, and so forms a basis.

(3) If two sets of vectors that belong to the same space are both a basis, then both sets must have the same number of vectors. That is, if $\boldsymbol{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ and $\boldsymbol{W} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_p\}$ form (two) basis in $\mathcal{V}$, it must be the case that $p = m$.

Proof. We begin with the fact that $\boldsymbol{V}$ is a basis. Then, any vector $\boldsymbol{w}_i$ necessarily is a linear combination of $\boldsymbol{V}$. We now use the property (2) inductively. We have that $\boldsymbol{w}_1 = \alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \alpha_3 \boldsymbol{v}_3 + \cdots + \alpha_m \boldsymbol{v}_m$. If $\alpha_1 \neq 0$, then $\{\boldsymbol{w}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_m\}$ forms a basis. Then, $\boldsymbol{w}_2 = \delta_1 \boldsymbol{w}_1 + \delta_2 \boldsymbol{v}_2 + \delta_3 \boldsymbol{v}_3 + \cdots + \delta_m \boldsymbol{v}_m$. If $\delta_2 \neq 0$, then $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_m\}$ forms a basis. And so on.
Suppose $p < m$. Following the described induction, we would have that $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_p, \boldsymbol{v}_{p+1}, \ldots, \boldsymbol{v}_m\}$ forms a basis. Then, $\boldsymbol{W} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_p\}$ could only span vectors in the subspace that place zero weights to $\{\boldsymbol{v}_{p+1}, \ldots, \boldsymbol{v}_m\}$, and not vectors in all of $\mathcal{V}$. This is a contraction. We discard $p < m$.
Consider now $p > m$. Following the induction, we would have that $\boldsymbol{W}^* = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$ forms a basis. Then, the vectors in $\{\boldsymbol{w}_{m+1}, \ldots, \boldsymbol{w}_p\}$ are *necessarily* linear combinations of the basis $\boldsymbol{W}^*$, and this would make $\boldsymbol{W}$ (that contains $\boldsymbol{W}^*$) linearly dependent. A contradiction. We also discard $p > m$, and conclude that $p = m$.

(4) Any linearly independent set of n vectors in $\mathbb{R}^n$ forms a basis of $\mathbb{R}^n$.

Proof. We use a similar argument than the one used to prove (3), i.e. using (2) inductively.
We begin with the canonical basis of $\mathbb{R}^n$, $\boldsymbol{H}_0 = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$, and sequentially generate vectors $\boldsymbol{v}_i$ such that $\boldsymbol{H}_1 = \{\boldsymbol{v}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \ldots, \boldsymbol{e}_n\}$, $\boldsymbol{H}_2 = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{e}_3, \ldots, \boldsymbol{e}_n\}$, $\boldsymbol{H}_3 = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{e}_n\}$ form basis of $\mathbb{R}^n$. The procedure stops when we reach the set $\boldsymbol{H}_n = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_n\}$. This set contains arbitrary combinations of $\boldsymbol{H}_0$. The only restriction is, of course, that, by definition of a basis, it must be linearly independent.

(5) A set of m vectors in $\mathbb{R}^n$ cannot span $\mathbb{R}^n$ if $m < n$.

Proof. Properties (3) and (4) indicate that a basis of $\mathbb{R}^n$ must have n vectors. A set of $m < n$ vectors in $\mathbb{R}^n$ can only span vector subspaces of $\mathbb{R}^n$ (of, at most, dimension m).

(6) Consider two sets of vectors from R^n, $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_p\}$ and $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$. Define also the matrices $\mathbf{U} = [\,\boldsymbol{u}_1\, \boldsymbol{u}_2 \cdots \boldsymbol{u}_p\,]$ (n rows, p columns), $\mathbf{V} = [\,\boldsymbol{v}_1\, \boldsymbol{v}_2 \cdots \boldsymbol{v}_m\,]$ (n rows, m columns) and $\mathbf{W} = [\,\mathbf{U}\,\mathbf{V}\,]$ (n rows, $p + m$ columns). Then, $\mathrm{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) \subset \mathrm{span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p)$ if and only if $\mathrm{rank}(\mathbf{U}) = \mathrm{rank}(\mathbf{W})$.
Moreover, $\mathrm{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) = \mathrm{span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p)$ if and only if $\mathrm{rank}(\mathbf{U}) = \mathrm{rank}(\mathbf{V}) = \mathrm{rank}(\mathbf{W})$.

Proof. This follows from the properties of the Gauss-Jordan elimination method.