Linear Independence

CE40282-1: Linear Algebra
Hamid R. Rabiee and Maryam Ramezani
Sharif University of Technology

Subspace

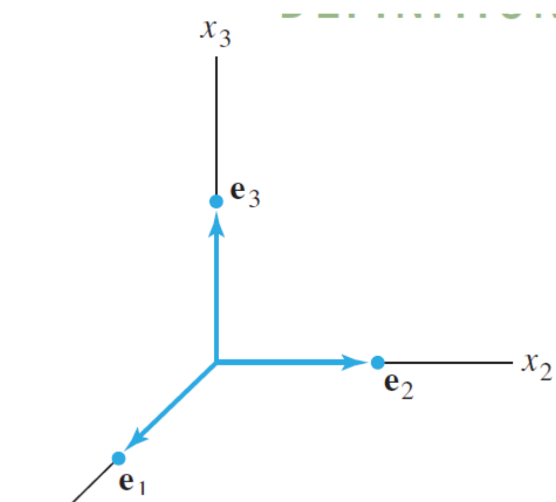
Definition 1.15. Let $(V, +, \cdot)$ be a vector space, and let $W \subset V$ be a subset. Then W is a *subspace* of V if the following properties are satisfied:

(1) The zero vector $\mathbf{0} \in V$ is in W.
(2) (Closed under $+$) For all $\mathbf{w}_1, \mathbf{w}_2 \in W$, we have $\mathbf{w}_1 + \mathbf{w}_2 \in W$.
(3) (Closed under $\cdot$) For all $\mathbf{w} \in W$ and $\lambda \in K$, we have $\lambda \cdot \mathbf{w} \in W$.

The axioms in the definition of a subspace ensure that the addition and scalar multiplication operations on V make sense as operations on W: if we add two vectors in W we get a vector in W, and if we scalar multiply a vector in W by a scalar, we get a vector in W.

Basis

- A set of n linearly independent n-vectors is called a basis
- A basis is the combination of span and independence: A set of vectors $\{v_1, \ldots, v_n\}$ forms a basis for some subspace of R^n if it
  - (1) spans that subspace
  - (2) is an independent set of vectors.



The standard basis for $\mathbb{R}^3$.

Basis

- Which are unique?
  - express a vector in terms of any particular basis
  - bases for R^2
  - bases with unit length for R^2

Coordinate Systems

- The main reason for selecting a basis for a subspace H; instead of merely a spanning set, is that each vector in H can be written in only one way as a linear combination of the basis vectors.

Suppose the set $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_p\}$ is a basis for a subspace H. For each $\mathbf{x}$ in H, the **coordinates of x relative to the basis** $\mathcal{B}$ are the weights $c_1, \dots, c_p$ such that $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_p\mathbf{b}_p$, and the vector in $\mathbb{R}^p$

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

is called the **coordinate vector of x (relative to** $\mathcal{B}$**)** or the $\mathcal{B}$**-coordinate vector of x**.[1]

Dimensions

- The number of elements in a vector
- In the vector [3 1 5], the element in the second dimension is 1, and the number 5 is in the third dimension.
- The dimensionality of a vector is the number of coordinate axes in which that vector exists.
- If a vector space is spanned by a finite number of vectors, it is said to be finite-dimensional. Otherwise it is infinite-dimensional.
- The number of vectors in a basis for a finite-dimensional vector space V is called the dimension of V and denoted dim V .

Coordinate axes

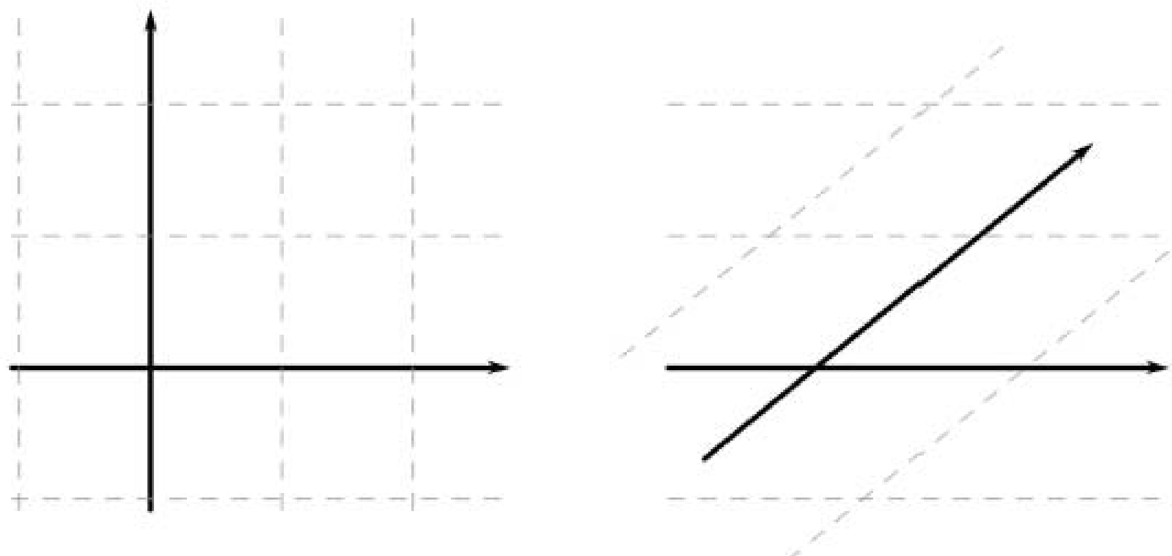
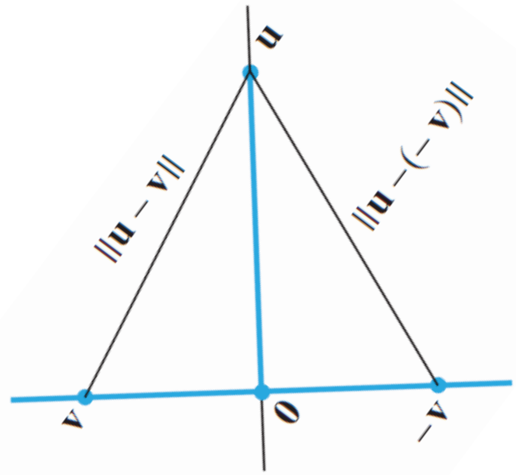


Figure 4.3: The familiar Cartesian plane (left) has orthogonal coordinate axes. However, axes in linear algebra are not constrained to be orthogonal (right), and non-orthogonal axes can be advantageous.

Orthogonal vectors

- Geometry:



- Algebra:
  - Two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are **orthogonal** (to each other) if $\mathbf{u} \cdot \mathbf{v} = 0$.
  - **The Pythagorean Theorem**
    Two vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$

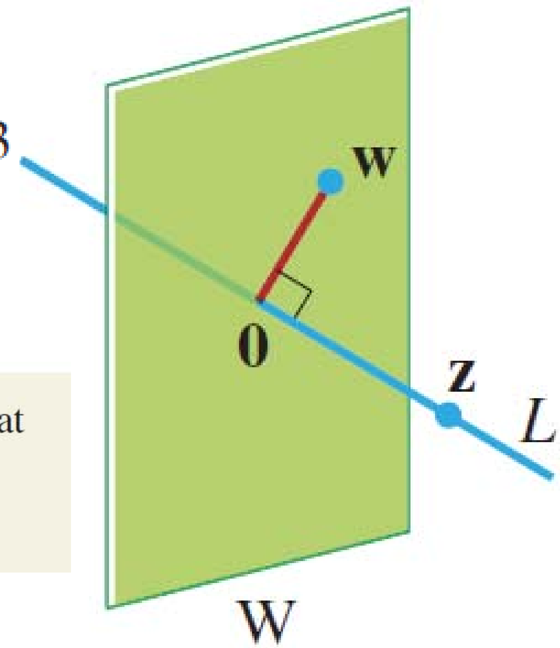
Orthogonal Complements

- If a vector z is orthogonal to every vector in a subspace W of R^n, then z is said to be orthogonal to W.
- The set of all vectors z that are orthogonal to W is called the orthogonal complement of W and is denoted by $w^{\perp}$

W be a plane through the origin in $\mathbb{R}^3$

$L = W^{\perp}$ and $W = L^{\perp}$

1. A vector $\mathbf{x}$ is in $W^{\perp}$ if and only if $\mathbf{x}$ is orthogonal to every vector in a set that spans W.
2. $W^{\perp}$ is a subspace of $\mathbb{R}^n$.



Orthogonal Sets

- A set of vectors $\{a_1, \ldots, a_k\}$ in R^n is orthogonal set if each pair of distinct vectors is orthogonal
- Theorem:
  - If $S = \{a_1, \ldots, a_k\}$ is an orthogonal set of nonzero vectors in R^n, then S is linearly independent and is a basis for the subspace spanned by S.
    - Proof?

Orthonormal vectors

- set of n-vectors $a_1, \ldots, a_k$ are *(mutually) orthogonal* if $a_i \perp a_j$ for $i \neq j$
- they are *normalized* if $\|a_i\| = 1$ for $i = 1, \ldots, k$
- they are *orthonormal* if both hold
- can be expressed using inner products as

$$a_i^T a_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

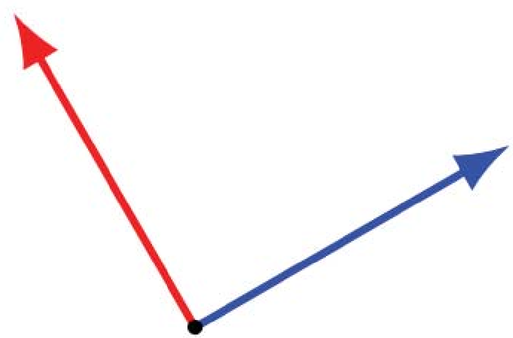
- orthonormal sets of vectors are linearly independent
- by independence-dimension inequality, must have $k \leq n$
- when $k = n$, $a_1, \ldots, a_n$ are an *orthonormal basis*

Examples of orthonormal bases

- standard unit n-vectors $e_1, \ldots, e_n$
- the 3-vectors

$$\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \quad \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

- the 2-vectors shown below



Linear combinations of orthonormal vectors

- A simple way to check if an n-vector y is a linear combination of the orthonormal vectors $a_1, \ldots, a_k$, if and only if:

$$y = (a_1^T y)a_1 + \cdots + (a_k^T y)a_k$$

- For orthogonal vectors $a_1, \ldots, a_k$:

$$y = c_1 a_1 + \cdots + c_k a_k$$

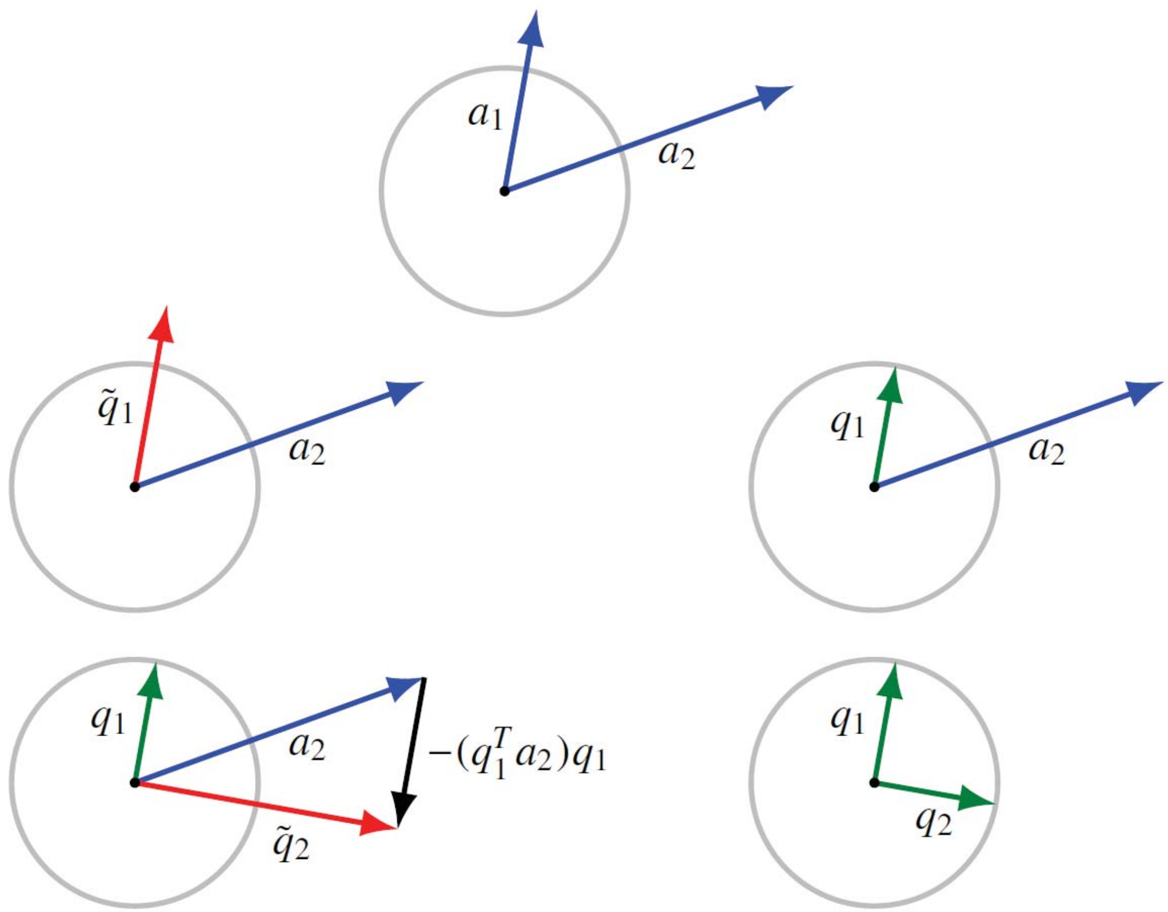
$$c_j = \frac{y.a_j}{a_j.a_j}$$

Example

$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad a_1 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \quad a_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad a_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Gram–Schmidt (orthogonalization) algorithm

- Find orthonormal basis for span$\{a_1, a_2, \dots, a_k\}$
- Geometry:



Gram–Schmidt (orthogonalization) algorithm

- Find orthonormal basis for span$\{a_1, a_2, \dots, a_k\}$
- Algebra:

1 $$q_1 = \frac{a_1}{\|a_1\|}$$

2 $$\tilde{q}_2 = a_2 - (q_1^T a_2) q_1 \quad \rightarrow q_2 = \frac{\tilde{q}_2}{\|\tilde{q}_2\|}$$

3 $$\tilde{q}_3 = a_3 - (q_1^T a_3) q_1 - (q_2^T a_3) q_2 \quad \rightarrow q_3 = \frac{\tilde{q}_3}{\|\tilde{q}_3\|}$$

$$\vdots$$

k $$\tilde{q}_k = a_k - (q_1^T a_k) q_1 - \dots - (q_{k-1}^T a_k) q_{k-1} \quad \rightarrow q_k = \frac{\tilde{q}_k}{\|\tilde{q}_k\|}$$

Gram–Schmidt (orthogonalization) algorithm

- Why $\{q_1, q_2, \dots, q_k\}$ is a orthonormal basis for span$\{a_1, a_2, \dots, a_k\}$?
  - $\{q_1, q_2, \dots, q_k\}$ are normalized.
  - $\{q_1, q_2, \dots, q_k\}$ is a orthogonal set
  - a_i is a linear combination of $\{q_1, q_2, \dots, q_i\}$

$$\Downarrow$$

span$\{q_1, q_2, \dots, q_k\}$=span$\{a_1, a_2, \dots, a_k\}$

- q_i is a linear combination of $\{a_1, a_2, \dots, a_i\}$

Gram–Schmidt (orthogonalization) algorithm

given n-vectors $a_1, \ldots, a_k$

for $i = 1, \ldots, k$

1. *Orthogonalization:* $\tilde{q}_i = a_i - (q_1^T a_i)q_1 - \cdots - (q_{i-1}^T a_i)q_{i-1}$
2. *Test for linear dependence:* if $\tilde{q}_i = 0$, quit
3. *Normalization:* $q_i = \tilde{q}_i / \|\tilde{q}_i\|$

- if G–S does not stop early (in step 2), $a_1, \ldots, a_k$ are linearly independent
- if G–S stops early in iteration $i = j$, then a_j is a linear combination of $a_1, \ldots, a_{j-1}$ (so $a_1, \ldots, a_k$ are linearly dependent)

$$\boldsymbol{a}_j = (\boldsymbol{q}_1^T \boldsymbol{a}_j)\boldsymbol{q}_1 + \cdots + (\boldsymbol{q}_{j-1}^T \boldsymbol{a}_j)\boldsymbol{q}_{j-1}$$

Complexity of Gram–Schmidt algorithm

given n-vectors $a_1, \ldots, a_k$

for $i = 1, \ldots, k$

1. *Orthogonalization:* $\tilde{q}_i = a_i - (q_1^T a_i)q_1 - \cdots - (q_{i-1}^T a_i)q_{i-1}$
2. *Test for linear dependence:* if $\tilde{q}_i = 0$, quit
3. *Normalization:* $q_i = \tilde{q}_i / \|\tilde{q}_i\|$

Reference

- Page 97 LINEAR ALGEBRA: Theory, Intuition, Code
- Page 213: David Cherney,
- Page 54: Linear Algebra and Optimization for Machine Learning