Engineering Decision Theory
EN5830

Chathuranga Weeraddana

December 2019

Lecture 5

Linearly Independence

Linearly Independent Set of Vectors

- a set of vectors $\{v_1, v_2, \ldots v_k\}$ is independent whenever

$$\underbrace{(\alpha_1 v_1 + \alpha_2 v_2 + \cdots \alpha_k v_k = 0)}_{A} \implies \underbrace{\alpha_i = 0 \text{ for all } i}_{B}$$

  an equivalent statement:

- no vector v_i can be expressed as a linear combination of the other vectors $v_1 \ldots, v_{i-1}, v_{i+1}, \ldots, v_k$

Linearly Independent Set of Vectors

- a set of vectors $\{v_1, v_2, \ldots v_k\}$ is dependent whenever

  $\text{not}(A \implies B)$, in other words, when $\Big((A) \text{ and } (\text{not } B) \Big)$

  an equivalent statement:

- at least one vector v_i can be expressed as a linear combination of the other vectors $v_1 \ldots, v_{i-1}, v_{i+1}, \ldots, v_k$

Linearly Independent Set of Vectors

- exercise: discuss the linear independence/dependence of following set $\mathcal{S}$:
  - $\mathcal{S} = \{v_1\}$, where $v_1 = (0, 0) \in \mathbb{R}^2$
  - $\mathcal{S} = \{v_1, v_2\}$, where $v_1 = (0, 0) \in \mathbb{R}^2$ and $v_2 = (1, 1) \in \mathbb{R}^2$
  - $\mathcal{S} = \{v_1, v_2\}$, where $v_1 = (2, 2) \in \mathbb{R}^2$ and $v_2 = (1, 1) \in \mathbb{R}^2$
  - $\mathcal{S} = \{v_1, v_2\}$, where $v_1 = (2, 1) \in \mathbb{R}^2$ and $v_2 = (1, 1) \in \mathbb{R}^2$
  - $\mathcal{S} = \{v_1, v_2, v_3\}$, where $v_1 = (1, 0)$, $v_2 = (0, 1)$, and $v_3 = (1, 1)$

Span

Span of a Set of Vectors

- a set of vectors is given: $\{a_1, a_2, \ldots, a_k\}$
- span$\{a_1, a_2, \ldots, a_k\}$ is the set of points that can be hit by any linear combination of the vectors in $\{a_1, a_2, \ldots, a_k\}$
- more specifically,

$$\text{span}\{a_1, a_2, \ldots, a_k\} = \left\{ \sum_{i=1}^{k} \beta_i a_i \;\middle|\; (\beta_1, \beta_2, \ldots, \beta_k) \in \mathbb{R}^k \right\}$$

Span of a Set of Vectors

- illustrate the idea graphically
- exercise:
  - Suppose $a = a_1 + 2a_k$. What can you say about span$\{a_1, a_2, \ldots, a_k\}$ and span$\{a_1, a_2, \ldots, a_k, a\}$? Are they same or different?
  - Suppose a cannot be written as a linear combination of vectors in $\{a_1, a_2, \ldots, a_k, a\}$. What can you say about span$\{a_1, a_2, \ldots, a_k\}$ and span$\{a_1, a_2, \ldots, a_k, a\}$? Are they same or different?

Subspaces

Subspace of $\mathbb{R}^n$

- a subset $\mathcal{V} \subseteq \mathbb{R}^n$ is called a subspace of $\mathbb{R}^n$ if $\mathcal{V}$ is closed under the operations of
  - vector addition, i.e.,

$$a, b \in \mathcal{V} \implies (a + b) \in \mathcal{V}$$

    and
  - scalar multiplication, i.e.,

$$a \in \mathcal{V} \implies \alpha a \in \mathcal{V} \text{ for all } \alpha \in \mathbb{R}$$

- exercise: is $0 \in \mathbb{R}^n$ an element of $\mathcal{V}$?
- draw some subspaces to trigger the intuition..

Basis of the Subspace $\mathcal{V}$

- given a subspace $\mathcal{V}$, any set of linearly independent set of vectors $\{a_1, a_2, \ldots, a_k\} \subset \mathcal{V}$ such that

$$\mathcal{V} = \text{span}\{a_1, a_2, \ldots, a_k\}$$

  is called a basis of $\mathcal{V}$

- all bases of a subspace $\mathcal{V}$ contain the same number of vectors

- we call it the dimension of $\mathcal{V}$ or dim $\mathcal{V}$

Basis of the Subspace $\mathcal{V}$

- exercise:
  - you are given a set of vectors $\{a_1, a_2, a_3\}$, where

$$a_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad a_2 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \qquad a_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

  - let $\mathcal{V} = \text{span}\{a_1, a_2, a_3\}$
  - what is $\mathcal{V}$?
  - give a basis for $\mathcal{V}$ by using vectors in $\{a_1, a_2, a_3\}$ itself
  - give a basis for $\mathcal{V}$ without using vectors in $\{a_1, a_2, a_3\}$ itself
  - what is dim $\mathcal{V}$?

Coordinates

- suppose we are given a basis $\{a_1, a_2, \ldots, a_k\}$ of $\mathcal{V}$ and a vector $a \in \mathcal{V}$
- thus, a can be expressed as

$$a = \beta_1 a_1 + \beta_2 a_2 + \cdots + \beta_k a_k \text{ for some } \beta \in \mathbb{R}^k$$

- the coefficients β_i, $i = 1, 2, \ldots, k$ are called the coordinates of a with respect to the basis $\{a_1, a_2, \ldots, a_k\}$

Coordinates

- exercise:
  - consider the basis $\{e_1, e_2\}$ of $\mathbb{R}^2$ and a vector $a = (1, 2) \in \mathbb{R}^2$
  - here $e_1 = (1, 0)$ and $e_2 = (0, 1)$
  - what are the coordinates of a with respect to the basis $\{e_1, e_2\}$?

Rank and Nullspace

Matrices

- a matrix is a rectangular array of numbers written between rectangular brackets, e.g.,

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 1 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 2}$$

- in general a matrix can have m rows and n columns, i.e., $A \in \mathbb{R}^{m\times n}$
- transpose of A is denoted $A^T \in \mathbb{R}^{n\times m}$
- note that $\mathbb{R}^{m\times n}$ is the set of real $m \times n$ matrices
- $\mathbb{R}^{n\times 1} = \mathbb{R}^n$ based on our notations

Matrices

- some useful compact representations of $A \in \mathbb{R}^{m \times n}$
- represent in terms of columns:

$$A = [a_1 \; a_2 \; \cdots a_n],$$

where $a_i \in \mathbb{R}^m$, $i = 1, 2, \ldots, n$

- represent in terms of rows:

$$A = \begin{bmatrix} \tilde{a}_1^T \\ \tilde{a}_2^T \\ \vdots \\ \tilde{a}_m^T \end{bmatrix},$$

where $\tilde{a}_i \in \mathbb{R}^n$, $i = 1, 2, \ldots, m$

Rank of a Matrix

- let the matrix $A \in \mathbb{R}^{m \times n}$ is given by

$$A = [a_1 \ a_2 \ \cdots a_n],$$

  where $a_i \in \mathbb{R}^m$, $i = 1, 2, \ldots, n$

- rank A = maximum number of linearly independent columns of A

- rank A = maximum number of linearly independent rows of A

- thus, rank $A \leq \min(m, n)$

Rank of a Matrix

- exercise:
  - determine the rank of following matrices:

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 2 \\ 3 & 6 & 0 \end{bmatrix}$$

  - Matlab can be used to easily determine the rank of a matrix, check
  - many algorithms to determine rank, e.g., GramSchmidt procedure

Rank of a Matrix

- rank of a matrix A is invariant under
  - multiplication of the columns of A by a nonzero scalar
  - interchange of columns
  - addition to a given column a linear combination of others
- check proposition 2.3 of ref-1

Nullspace of a Matrix

- the nullspace of $A \in \mathbb{R}^{m \times n}$ is given by

$$\mathcal{N}(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}$$

- $\mathcal{N}(A)$ is the set of vectors mapped to 0 when multiplied by A
- $\mathcal{N}(A)$ is the set of vectors orthogonal to rows of A
- $\mathcal{N}(A)$ gives ambiguity in x given $y = Ax$
  - if $y = Ax$ and $z \in \mathcal{N}(A)$, then $y = A(x + z)$

Zero Nullspace

- 0 is the only element of $\mathcal{N}(A)$, i.e., $\mathcal{N}(A) = \{0\}$
- $\mathcal{N}(A) = \{0\} \iff$
  - x can always be uniquely determined from $y = Ax$
  - A is one-to-one, i.e., different x's map to different y's
  - columns of A are independent (hence, a basis for span$\{a_1, \ldots, a_n\}$)
  - det$(A^T A) \neq 0$
- all of the above equivalences can be proved, some obvious, some not

Range of a Matrix

- the range of $A \in \mathbb{R}^{m \times n}$ is given by

$$\mathcal{R}(A) = \{Ax \mid x \in \mathbb{R}^n\}$$

- note that $\mathcal{R}(A) \subseteq \mathbb{R}^m$
- $\mathcal{R}(A)$ is the set of vectors that can be hit by mapping $y = Ax$
- $\mathcal{R}(A) = \text{span}\{a_1, \ldots, a_n\}$
- $\mathcal{R}(A)$ is the set of vectors y for which $Ax = y$ has a solution

Onto Matrix

- $\mathcal{R}(A)$ is identical to R^m, i.e., $\mathcal{R}(A) = \mathbb{R}^m$, A is called onto
- $\mathcal{R}(A) = \mathbb{R}^m \iff$
  - $Ax = y$ can be solved in x for any y
  - columns of A spans $\mathbb{R}^m$
  - rows of A are independent
  - $\det(AA^T) \neq 0$
  - $\mathcal{N}(A^T) = \{0\}$
- all of the above equivalences can be proved, some obvious, some not

Square Matrices

- number of rows is equal to the number of columns
- for such a square matrices A we define the determinant of A, denoted $|A|$ or $\det A$
- if $|A| \neq 0$, we can also define the inverse of A, denoted A^{-1}
- read pages 14-15 of [Chong-2013], (more details chapter 6 of]Mayer-2000])
- special square matrices: e.g., identity matrix, diagonal matrices

Linear Equations and Their Solution

Linear Equations

- suppose we are given m equations in n unknowns of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

- define $a_i = (a_{1i}, a_{2i}, \ldots, a_{mi}) \in \mathbb{R}^m$, $i = 1, 2, \ldots, n$

- then the above set of equations can also be given by

$$a_1x_1 + a_2x_2 + \cdots a_nx_n = b,$$

where $b = (b_1, b_2, \ldots, b_m) \in \mathbb{R}^m$

Linear Equations

- we can further compactly represent the above equations as follows

$$[a_1 \ a_2 \ \cdots a_n] \ x = b,$$

where $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$

- if we let $A = [a_1 \ a_2 \ \cdots a_n]$, we get the system of equations

$$A \ x = b$$

- it is recommended to be familiar with block matrix operations!

Solution of a Linear System

- there can be
  - many solutions
  - unique solution
  - no solution
- exemplify the above 3 case !

Solution of a Linear System

- let $[A\ b]$ denote the matrix by concatenating $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$
- The linear system $Ax = b$ has a solution if and only if

$$\text{rank } A = \text{rank } [A\ b]$$

- see Theorem 2.1 of ref-1 for a proof!
- the above suggests that

$$\dim\big(\text{span}\{a_1, a_2, \ldots, a_n\}\big) = \dim\big(\text{span}\{a_1, a_2, \ldots, a_n, b\}\big)$$

- i.e., b can be written as a linear combination of the vectors $a_1, a_2, \ldots, a_n$

Solution of a Linear System

- More specifically, $Ax = b$, $A \in \mathbb{R}^{m \times n}$
- has no solution:

$$\text{rank}A < \text{rank}[A\ b]$$

- has solution:
  - unique: rank A = rank $[A\ b] = n$
  - many: rank A = rank $[A\ b] < n$
- graphically illustrate the 3 cases above