EE2211 Pre-Tutorial 3

Dr Feng LIN
feng_lin@nus.edu.sg

Agenda

- Recap
- Self-learning
- Tutorial 3

Today's Attendance

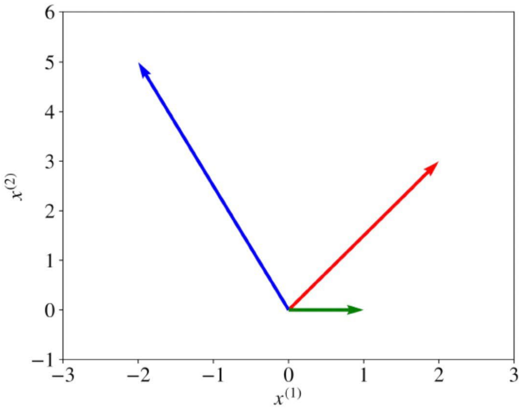
Recap

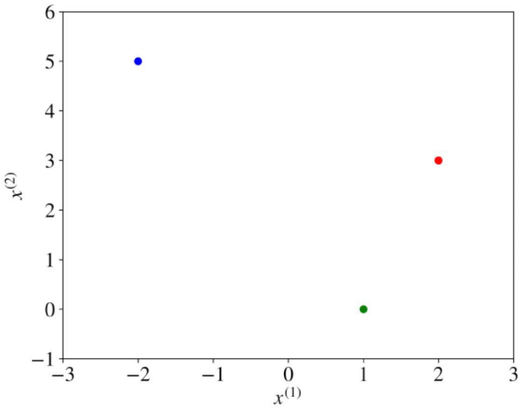
- Linear Algebra
  - Vectors
  - Matrices
  - Linear Equations
- Probability
  - Axioms of Probability
  - Random Variable
  - Basic rules
  - Bayes' rule

Vectors

- A **vector** is an ordered list of scalar values
  - Denoted by a bold character, e.g. x or a
- **Vectors** can be visualized as, in a multi-dimensional space,
  - arrows that point to some directions, or e.g. Velocity vector
  - points e.g. color points

Illustrations of three two-dimensional vectors, $\mathbf{a} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$, and $\mathbf{c} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$





Matrices

- A matrix is a rectangular array of numbers arranged in rows and columns
  - ➢ Denoted with bold capital letters, such as X or W
  - ➢ An example of a matrix with two rows and three columns:

$$\mathbf{X} = \begin{bmatrix} 2 & 4 & -3 \\ 21 & -6 & -1 \end{bmatrix}$$

e.g. image frame

- Note:
  - ➢ For elements in matrix X, we shall use the indexing $x_{1,1}$ where the first and second indices indicate the row and the column position.
  - ➢ Usually, for input data, rows represent samples and columns represent features

Linear Equation

Linear dependence and independence

- A collection of d-vector $\mathbf{x}_1, \cdots, \mathbf{x}_m$ (with $m \geq 1$) is called linearly dependent if

$$\beta_1 \mathbf{x}_1 + \cdots + \beta_m \mathbf{x}_m = 0$$

Holds for some $\beta_1, \cdots, \beta_m$ that are not all zeros.

- A collection of d-vector $\mathbf{x}_1, \cdots, \mathbf{x}_m$ (with $m \geq 1$) is called linearly independent if

$$\beta_1 \mathbf{x}_1 + \cdots + \beta_m \mathbf{x}_m = 0$$

only hold for $\beta_1 = \cdots = \beta_m = 0$.

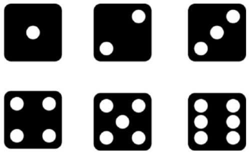
Note: If all rows or columns of a square matrix **X** are linearly **independent**, then **X** is **invertible**.

Probability

- We describe a *random experiment* by describing its procedure and observations of its *outcomes*.

Outcome

- Outcomes are mutual exclusive in the sense that only one outcome occurs in a specific trial of the random experiment.
  - This also means an outcome is not decomposable.
  - All unique outcomes form a *sample space.*



Sample space

- A subset of sample space S, denoted as A, is an *event* in a random experiment $A \subset S$, that is meaningful to an application.
  - Example of an event: faces with numbers no greater than 3

Event

Axioms of Probability

- Assume events $A \subseteq S$ and $B \subseteq S$.

1. **[Nonnegativity]** $Pr(A) \geq 0$ for every event A.
2. **[Normalization]** $Pr(S) = 1$.
3. **[Additivity]** If $A \cap B = \emptyset$, then $Pr(A \cup B) = Pr(A) + Pr(B)$.

   otherwise $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$

   It can be extended to more events:

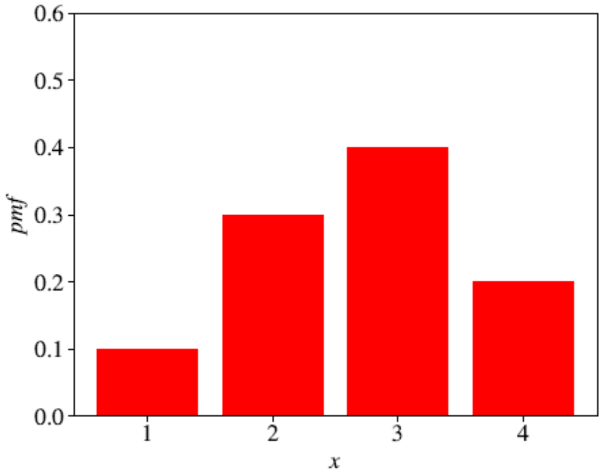
   If $A_1, A_2, \ldots$ are disjoint, then $Pr(\cup_{k=1}^{\infty} A_k) = \Sigma_{k=1}^{\infty} Pr(A_k)$.

Random Variable

- A random variable, usually written as an italic capital letter, like X, is a variable whose possible values are numerical outcomes of a random event.
- There are two types of random variables: discrete and continuous.

Discrete Random Variable

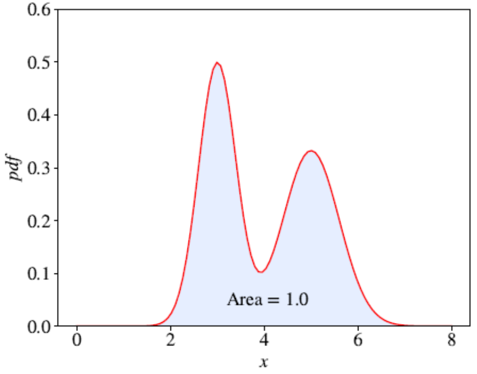
- A discrete random variable (DRV) takes on only a countable number of distinct values such as red, orange, blue or 1, 2, 3.
- The probability distribution of a discrete random variable is described by a list of probabilities associated with each of its possible values.
- This list of probabilities is called a probability mass function (pmf).
  - ➢ Like a histogram, except that here the probabilities sum to 1



A probability mass function

Continuous Random Variable

- A continuous random variable (CRV) takes an infinite number of possible values in some interval.
  - ➢ Examples include height, weight, and time.
  - ➢ The number of values of a continuous random variable X is infinite, the probability $\Pr(X = c)$ for any c is *0.*
  - ➢ Therefore, instead of the list of probabilities, the probability distribution of a CRV (a continuous probability distribution) is described by a **probability density function** (pdf).
  - ➢ The pdf is a function whose range is nonnegative and the area under the curve is equal to 1.



A probability density function

Two Basic Rules

- Sum Rule

$$\Pr(X = x) = \sum_{Y} \Pr(X = x, Y = y_i)$$

- Product Rule

$$\Pr(X = x, Y = y) = \Pr(Y = y | X = x)\, P(X = x)$$

Bayes' Rule

- The conditional probability $\Pr(Y = y|X = x)$ is the probability of the random variable Y to have a specific value y, given that another random variable X has a specific value of x.

- The **Bayes' Rule** (also known as the **Bayes' Theorem**):

$$\underbrace{\Pr(Y = y|X = x)}_{\text{posterior}} = \frac{\overbrace{\Pr(X = x|Y = y)}^{\text{likelihood}}\,\overbrace{\Pr(Y = y)}^{\text{prior}}}{\underbrace{\Pr(X = x)}_{\text{evidence}}}$$

Example of Bayes' Rule

- A guy was tested positive for a rare disease!
  - In the general population only 0.1% has this disease
- The test is quite accurate
  - It can detect 99% of subjects with the disease
  - It only mis-identifies 2% of subjects without this disease

How worried should this guy be?

Example of Bayes' Rule

	Simple space (1,000 people)	Test Positive	Probability
Disease	1	1	~1/21
No Disease	999	20	~20/21

Example of Bayes' Rule

$$P(H|E) = \frac{P(E|H)P(H)}{P(E|H)P(H) + P(E|\overline{H})P(\overline{H})}$$

$P(H)$: Chance of having this disease = 0.001 (Prior probability)

$P(\overline{H}) = 0.999$

$P(E|H)$: Sensitivity = 0.99 (Likelihood)

$P(E|\overline{H})$: false positive rate = 0.02

$$P(H|E) = \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.02 \times 0.999} = 0.0472$$

Example of Bayes' Rule

If tested positive the 2nd time

$$P(H|E) = \frac{P(E|H)P(H)}{P(E|H)P(H) + P(E|\overline{H})P(\overline{H})}$$

$P(H) = 0.0472$ (Posterior probability: tested positive the 1st time)

$P(\overline{H}) = 0.9528$

$P(E|H)$: Sensitivity = 0.99

$P(E|\overline{H})$: False positive rate = 0.02

$$P(H|E) = \frac{0.99 \times 0.0472}{0.99 \times 0.0472 + 0.02 \times 0.9528} = 0.7103$$

THANK YOU