Probabilistic Robotics and Inference Algorithms

Chandra Gummaluru

November 20, 2019

Contents

Introduction

Any real robot must deal with two major problems:

1. the physical world is inherently unpredictable and;
2. robotic sensors are inherently noisy.

For this reason, many models of robotic systems are probabilistic.

Our goal is to use a probabilistic model of our robot and its environment to make inferences about said environment based on knowledge we gain from interacting with it.

Naturally, we will be relying heavily on probability theory.

1 Set Theory

Much of probability theory involves the manipulation of sets.

A **set**, S is simply a collection of items. Each item, s, is referred to as an **element** of S, which we denote as $s \in S$.

1.1 Sets Operations

1.1.1 Complement

The **complement** of a set A, denoted $A^{\complement}$, is the set of elements not contained in A. That is,

$$a \in A \Leftrightarrow a \notin A^{\complement}$$

1.1.2 Union

The **union** of two sets, A,B, denoted $A \cup B$, is the set of elements contained in either set A or B. It follows that,

$$s \in A, B \Rightarrow s \in (A \cup B)$$

In general, for n sets, $S_1, \ldots, S_n$,

$$S_1 \cup \ldots \cup S_n = \bigcup_{k=1}^{n} S_k$$

1.1.3 Intersection

The **intersection** of two sets, A, B, denoted, $A \cap B$, is the set of elements contained in both sets. That is,

$$s \in (A \cap B) \Leftrightarrow s \in A, s \in B$$

In general, for n sets, $S_1, \ldots, S_n$,

$$S_1 \cap \ldots \cap S_n = \bigcap_{k=1}^{n} S_k$$

Two sets are said to be **disjoint** if their intersection is the empty set, i.e. they share no elements.

1.1.4 Difference

The **difference** or **relative complement** between two sets, A, B, denoted by $A - B$, is the elements of A not contained in B. That is,

$$s \in (A - B) \Rightarrow s \in A, s \notin B$$

1.2 Theorems of Set Theory

1.2.1 DeMorgan's Laws

For any two finite sets, A, B,

(i) $(A \cup B)^{\mathsf{C}} = A^{\mathsf{C}} \cap B^{\mathsf{C}}$ (Union Law)

Proof:
Suppose $s \in (A \cup B)^{\mathsf{C}}$. Then, $s \notin (A \cup B)$.
$\Rightarrow s \notin A, s \notin B$
$\Rightarrow s \in A^{\mathsf{C}}, s \in B^{\mathsf{C}}$
$\Rightarrow s \in (A^{\mathsf{C}} \cap B^{\mathsf{C}})$
$\therefore (A \cup B)^{\mathsf{C}} \subset (A^{\mathsf{C}} \cap B^{\mathsf{C}})$
Now, suppose $s \in (A^{\mathsf{C}} \cap B^{\mathsf{C}})$. Then, $s \in A^{\mathsf{C}}, s \in B^{\mathsf{C}}$.
$\Rightarrow s \notin A, s \notin B$
$\Rightarrow s \notin (A \cup B)$
$\Rightarrow s \in (A \cup B)^{\mathsf{C}}$
$\blacksquare$

(ii) $(A \cap B)^{\mathsf{C}} = A^{\mathsf{C}} \cup B^{\mathsf{C}}$ (Intersection Law)

Proof:
Suppose $s \notin (A \cap B)^{\mathsf{C}}$. Then, $s \in (A \cap B)$.
$\Rightarrow s \in A, s \in B$
$\Rightarrow s \notin A^{\mathsf{C}}, s \notin B^{\mathsf{C}}$
$\Rightarrow s \notin (A^{\mathsf{C}} \cup B^{\mathsf{C}})$
$\therefore (A \cup B)^{\mathsf{C}} \subset (A^{\mathsf{C}} \cap B^{\mathsf{C}})$
Now, suppose $s \in (A^{\mathsf{C}} \cap B^{\mathsf{C}})$. Then, $s \in A^{\mathsf{C}}, s \in B^{\mathsf{C}}$.
$\Rightarrow s \notin A, s \notin B$
$\Rightarrow s \notin (A \cup B)$
$\Rightarrow s \in (A \cup B)^{\mathsf{C}}$
$\blacksquare$

2 Probability Theory

With a basic understanding of set theory, we can define what it means to operate in a probabilistic space.

2.1 Probability Spaces

A **probability space** is a mathematical construct that models non-deterministic processes. It consists of three parts:

- a set, $\mathbb{S}$, of possible outcomes, called the **sample space**,
- a set, $\mathcal{F}$, of events, where an **event** is a set of zero or more outcomes, i.e., a subset of $\mathbb{S}$, and
- a **probability assignment function**, $P : \mathbb{S} \to [0, 1]$, where $P(o), o \in \mathbb{S}$, is the relative frequency of o occurring.

We say an event $A \in \mathcal{F}$ occurs if any outcome, $o \in A$ occurs. For any event, $A \in \mathcal{F}$, we say A

2.1.1 The Axioms of Probability

Any probability assignment function, P, must satisfy a few basic axioms:

- **The Axiom of Non-Negativity**: The probability of any outcome, $o \in \mathbb{S}$, occurring, is strictly non-negative, that is, $P[o] \geq 0$.
- **The Axiom of Conjunction** The probability of any event, A is

$$P[A] = \sum_{\forall o \in A} P[o].$$

- **The Axiom of Certainty**: Of all the possible outcomes in a sample space, one must occur, hence, $P[\mathbb{S}] = 1$.

Using these axioms we may derive more useful laws.

2.2 Probabilities of Events

2.2.1 Probability of Complements

The probability of the complement of an event is the complement of the probability. That is,

$$P\left[A^{\mathsf{C}}\right] = 1 - P[A]$$

Proof Since an event and its complement are mutually exclusive, definition, it follows that $A \cap A^{\mathsf{C}} = \emptyset$.

$$\therefore P\left[A \cup A^{\mathsf{C}}\right] = P[A] + P\left[A^{\mathsf{C}}\right]$$

However, since $A \cup A^{\mathsf{C}} = S$, it follows that

$$P[A] + P\left[A^{\mathsf{C}}\right] = P[S]$$

But, $P[S] = 1$.

$$\therefore P[A] + P\left[A^{\mathsf{C}}\right] = 1$$

$$\Rightarrow P\left[A^{\mathsf{C}}\right] = 1 - P[A]$$

■

2.2.2 Maximum Probability of an Event

The probability of any event cannot be larger than the probability of the certain event. That is,

$$P[A] \leq 1$$

Proof

$$P[A] = 1 - P\left[A^{\mathsf{C}}\right]$$

But, since $P\left[A^{\mathsf{C}}\right] \geq 0$,

$$P[A] \leq 1$$

■

2.2.3 Probability of the Impossible Event

The probability of the impossible event is zero. That is, $P[\emptyset] = 0$.

Proof: Let $A = S$. Then, $A^{\mathsf{C}} = \emptyset$. Now,

$$P[\emptyset] = P\left[A^{\mathsf{C}}\right] = 1 - P[A]$$

But $P[A] = P[S] = 1$.
$\therefore P[\emptyset] = 1 - 1$
$= 0$
■

Note that an event, A, with $P[A] = 0$ can still occur.

2.2.4 Probability of Mutually Exclusive Event

The probability of any one of a set of N mutually exclusive events, $\{A_i : i \leq N\}$, occurring is the sum of the probability of each event occurring. That is,

$$P\left[\bigcup_{k=1}^{N} A_k\right] = \sum_{k=1}^{N} P\left[A_k\right], n \geq 2$$

Proof:
For $N = 2$, we have $P\left[A_1 \cup A_2\right] = P\left[A_1\right] + P\left[A_2\right]$, which is the axiom of total probability mass.

Assume the result holds for $N = n$. That is,

$$P\left[\bigcup_{k=1}^{n} A_k\right] = \sum_{k=1}^{n} P\left[A_k\right]$$

Now,

$$\begin{aligned}
\sum_{k=1}^{n+1} P\left[A_k\right] &= \sum_{k=1}^{n} P\left[A_k\right] + P\left[A_{n+1}\right] \\
&= P\left[\bigcup_{k=1}^{n} A_k\right] + P\left[A_{n+1}\right] \\
&= P\left[\left\{\bigcup_{k=1}^{n} A_k\right\} \cup A_{n+1}\right] \\
&= P\left[\bigcup_{k=1}^{n+1} A_k\right]
\end{aligned}$$

Therefore, the result also holds for $N = n + 1$. By induction, the result must hold $\forall N \geq 2$.

$\blacksquare$

2.2.5 Probability of Non-Mutually Exclusive Events

Let A and B be two events, not necessarily mutually exclusive. Then,

$$P[A \cup B] = P[A] + P[B] - P[A \cap B]$$

Proof:
Decomposing the union of the two events into the sum of disjoint events,

$$P[A \cup B] = P\left[A \cap B^{\complement}\right] + P\left[A^{\complement} \cap B\right] + P[A \cap B]$$

Then,
$P[A] = P\left[A \cap B^{\complement}\right] + P[A \cap B]$
$\Rightarrow P\left[A \cap B^{\complement}\right] = P[A] - P[A \cap B]$
Similarly,
$P[B] = P\left[A^{\complement} \cap B\right] + P[A \cap B]$
$\Rightarrow P\left[A^{\complement} \cap B\right] = P[B] - P[A \cap B]$

$\therefore P[A \cup B] = (P[A] - P[A \cap B]) + (P[B] - P[A \cap B]) + P[A \cap B]$
$= P[A] + P[B] - P[A \cap B]$
■

2.2.6 Probability of Subsets

If $A \subset B$, then $P[A] \leq P[B]$.

Proof:
$A \subset B \Rightarrow B = A \cup \left(A^{\complement} \cap B\right)$
$\therefore P[B] = P[A] + P\left[A^{\complement} \cap B\right]$
$\geq P[A] \because P\left[A^{\complement} \cap B\right] \geq 0$
$\therefore P[A] \leq P[B]$
■

2.3 Conditional Probabilities of Events

In many cases, knowledge that an event has occurred influences the probability of some other event occurring.

2.3.1 The Definition of Conditional Probability

Consider two events, A and B, where $P[A] \neq 0$. By definition,

$$P[B] = \sum_{\forall o \in B} P[o].$$

Let $P[B|A]$ denote the probability of B, assuming A has occurred. In this case, we should only consider the outcomes of B, consistent with A, i.e., $A \cap B$.

$$\therefore P[B|A] = \sum_{\forall o \in A \cap B} P[o|A].$$

However, in general, $P[o|A] \neq P[o]$ for any $o \in A \cap B$. This is because $P[o]$ gives the relative frequency across $\mathbb{S}$, but we now only consider the outcomes of $\mathbb{S}$ consistent with A.

Thus,

$$P[o|A] = \frac{P[o]}{P[A]}$$

$$\begin{aligned}
\Rightarrow P[B|A] &= \sum_{\forall o \in A \cap B} \frac{P[o]}{P[A]} \\
&= \frac{\sum_{\forall o \in A \cap B} P[o]}{P[A]} \\
&= \frac{P[A \cap B]}{P[A]}.
\end{aligned}$$

Conditional probability assignment functions are valid probability assignment functions, i.e., they satisfy all the axioms of probability.

2.3.2 Independence of Events

The concept of conditional probabilities allows us to define a notion of independence between events.

Two events, A and B are **independent** if the occurrence of one does not influence the probability of the other, that is,

$$P[B|A] = P[B].$$

If A and B are independent, it can be shown that $P[A \cap B] = P[A]P[B]$.

We say A and B are **conditionally independent** on C if

$$P[B|A, C] = P[B|C].$$

The implication of this is that given knowledge of C, knowledge of A does not influence the probability of B. This notion of knowledge influencing probabilities is crucial to inferencing.

2.4 Random Variables

The outcomes in a sample space can refer to anything, i.e., $\mathbb{S}$ is a set of undefined objects.

Rather than manipulating these objects directly, a mapping, $X : \mathbb{S} \to \operatorname{dom} X$, from a sample space, $\mathbb{S}$ to a measurable space, $\operatorname{dom} X$.

We refer to X as a **random variable** and for any $D \subseteq \operatorname{dom} X$, the probability that X takes on some value in D, is

$$P[X \in D] = P[\{o \in \mathbb{S} : X(o) \in D\}].$$

2.4.1 Probability Density Functions

For any random variable, X, we will assume there exists a **probability density function**, $p_X : \operatorname{dom} X \to [0, 1]$, so that if $D \subseteq \operatorname{dom} X$,

$$\int_D p_X(x')dx' = P[X \in D].$$

Note that

$$\int_{\operatorname{dom} X} p_X(x')dx = P[X \in \operatorname{dom} X] = 1,$$

since X must take on some value in $\operatorname{dom} X$.

References

[1] Garcia, Alberto. *Probability, Statistics and Random Processes for Electrical Engineering.* Pearson/Prentice Hall, Reading, Massachusetts, 1993.

[2] Thrun, Sebastian. Burgard Wolfram. Dieter Fox. *Probabilistic Robotics*, Reading, 1999.