Robotic Localization and Mapping - Miniset 1
Basic Probability & Homogeneous Transformations

Prof. Joshua Mangelson
Brigham Young University

September 14, 2021

Probability Review

PR-Q1. For each of the following functions specify whether the function is a valid probability mass functions (PMF), a valid probability density function (PDF), or not a valid PMF or PDF. For each function that is not a valid PMF/PDF, specify the axiom that is violated.

a)

$$P(X=x)=\begin{cases} 0.2 & \text{if } x=1 \\ 0.4 & \text{if } x=2 \\ 0.5 & \text{if } x=3 \\ 0 & \text{otherwise.} \end{cases}$$

b)

$$P(X=x)=\begin{cases} 0.3 & \text{if } x=1 \\ 0.4 & \text{if } x=2 \\ 0.3 & \text{if } x=3 \\ 0 & \text{otherwise.} \end{cases}$$

c)

$$P(X=x)=\begin{cases} 0.8 & \text{if } x=1 \\ 0.3 & \text{if } x=2 \\ -0.1 & \text{if } x=3 \\ 0 & \text{otherwise.} \end{cases}$$

d)

$$f_X(x)=\begin{cases} 1 & \text{if } 0 \le x \le 0.5 \\ 2 & \text{if } 0.5 < x \le 0.75 \\ 0 & \text{otherwise.} \end{cases}$$

e)
Any $f_X(x)$ where, given

$$a < b < c,$$

$$\int_a^c f_X(x) \neq \int_a^b f_X(x) + \int_b^c f_X(x)$$

f)

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Solution

a) Invalid PMF - Axiom 2 is violated, probabilities do not sum to one.

b) Valid PMF

c) Invalid PMF - Axiom 1 is violated, probabilites are not between 0 and 1

d) Valid PDF

e) Invalid PDF - Axiom 3 is violated.

f) Valid PDF - This is the PDF of a scalar Gaussian random variable.

PR-Q2. A joint PMF/PDF encodes the joint probability that multiple random variables will take on certain values at the same time. Figure 1 encodes one possible joint PMF for two random variables Study (representing the amount of time you spend studying) and Grade (representing your grade). Lets pretend that this data pmf was estimated by collecting and analyzing data from an arbitrary math class.

	Study = 2hr	Study = 4 hr	Study = 6 hr
Grade = A	0.04	0.1	0.3
Grade = B	0.07	0.18	0.1
Grade = C	0.15	0.05	0.01

Figure 1: Joint probability mass function for the random variables study and grade.

Using this joint probability distribution, answer the following questions:

a) If we randomly pick a student, what is the likelihood (or probability) that the chosen student got an A and studied for 6 hours?

b) If we randomly pick another student, what is the likelihood (or probability) that the chosen student got a C and studied for 6 hours?

c) Now lets say we are only interested in the student's grade. We can summarize information from the joint distribution so that we can focus on the variable we care about. This is called **marginalization**. In doing so, we remove one of the random variables from the distribution to get a new distribution that summarizes or includes the information about the removed random variable. We call this **marginalizing out** a random variable. Marginalize out Study so that we get a distribution over Grade alone that tells us what grade we are likely to achieve (while taking into account all possible study times). State both the marginal distribution and the "marginalization" equation from the slides that best describes what you did to obtain it. What is the marginal or overall probability of getting an A?

d) We can also use a joint pmf to look at the likelihood of one random variable given another. This is called **conditioning**. State the discrete version of the definition of conditional probability, then use it to find the distribution over possible grades given that the student studied for 6 hours. Write out this conditional probability distribution. What is the probability of getting an A given that the student studied for 6 hours?

Solution

a) $p(Study = 6, Grade = A) = 0.3$

b) $p(Study = 6, Grade = C) = 0.01$

c)

$$p(Grade = g) = \begin{cases} 0.44 & \text{if } g = A \\ 0.35 & \text{if } g = B \\ 0.21 & \text{if } g = C \\ 0 & \text{otherwise.} \end{cases}$$

We get this by marginalizing out study time according to the following equation:

$$p(Grade = g) = \sum_{s} p(Study = s, Grade = g).$$

We can then use the resulting PMF to get:

$$p(Grade = A) = 0.44$$

d) The definition of conditional probability is:

$$p(X = x | Y = y) = \frac{p(X = x, Y = y)}{p(Y = y)}.$$

We can apply this to get the following:

$$p(Grade = g | Study = s) = \begin{cases} \begin{cases} \frac{0.04}{0.26} \text{ if } g = A \\ \frac{0.07}{0.26} \text{ if } g = B \\ \frac{0.15}{0.26} \text{ if } g = C \end{cases} & \text{if } s = 2 \\ \begin{cases} \frac{0.1}{0.33} \text{ if } g = A \\ \frac{0.18}{0.33} \text{ if } g = B \\ \frac{0.05}{0.33} \text{ if } g = C \end{cases} & \text{if } s = 4 \\ \begin{cases} \frac{0.3}{0.41} \text{ if } g = A \\ \frac{0.1}{0.41} \text{ if } g = B \\ \frac{0.01}{0.41} \text{ if } g = C \end{cases} & \text{if } s = 6 \\ \quad 0 & \text{otherwise.} \end{cases}$$

Given this, we can see that:

$$p(Grade = A | Study = 6) = 0.7317$$

.

PR-Q3. State the definition of independence for two random events. Now state the definition of conditional independence. Show (by example) that independence does not imply conditional independence.

Solution

Def. Two random events X and Y are called independent if:

$$p(X \cap Y) = p(X)p(Y).$$

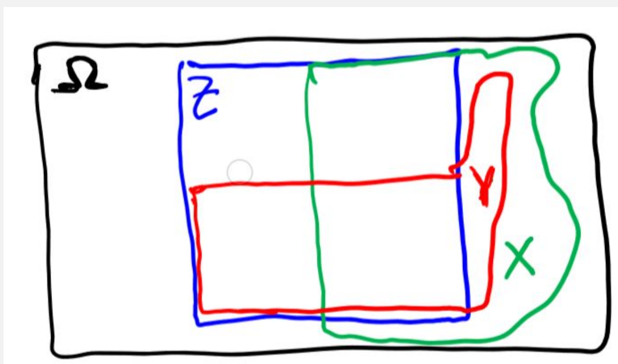
Def. Two random events X and Y are called conditionally independent given event Z if:

$$p(X \cap Y|Z = z) = p(X|Z)p(Y|Z).$$

Conditional independence does NOT imply independence. This can be seen from the following example:
The definition of conditional independence combined with the definition of conditional probabilities tells us that:

$$p(X|Z) = \frac{p(X \cap Z)}{p(Z)}, \quad p(Y|Z) = \frac{p(Y \cap Z)}{p(Z)}.$$

Assume we randomly draw a point from the box of possible outcomes Ω, shown below:



We can see that $p(X|Z) = p(X|Y \cap Z) = 0.5$ which implies that, given that we know the point was drawn in Z, additionally knowing the point is contained in Y does not change the likelihood that the point was also in X. This implies conditional independence.
However, if we ignore Z, independence does not hold for X and Y alone.

PR-Q4. Bayes rule/formula is very useful tool that enables us to rewrite a conditional probability in terms of other conditional and marginal distributions that are sometimes easier to work with. A common form of the Bayes rule is:

$$P(x|y) = \frac{P(y|x)P(x)}{\sum_x P(y|x)P(x)}.$$

Derive the Bayes rule from the definition of a conditional probabilities and the law of total probability.

Solution

From the definitions of conditional independence:

$$p(x, y) = p(y|x)p(x) = p(x|y)p(y).$$

Rearranging terms gives us:

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}.$$

Plugging in the law of total probability for $p(y)$ gives us:

$$P(x|y) = \frac{P(y|x)P(x)}{\sum_x P(y|x)P(x)}.$$

Rotation Matrices

RM-Q1. Show[1] that the dot product of two free vectors is the same regardless of the coordinate frame in which they are represented. Note that $\mathbf{v}_1 \cdot \mathbf{v}_2 = \mathbf{v}_1^\top \mathbf{v}_2$.

Solution

Showing this is equivalent to showing:

$$R\mathbf{v}_1 \cdot R\mathbf{v}_2 = \mathbf{v}_1 \cdot \mathbf{v}_2.$$

Using the definition above for dot product as well as the fact that the inverse of a rotation matrix is its transpose,

$$\begin{aligned} R\mathbf{v}_1 \cdot R\mathbf{v}_2 &= (R\mathbf{v}_1)^\top (R\mathbf{v}_2) \\ &= \mathbf{v}_1^\top R^\top R \mathbf{v}_2 \\ &= \mathbf{v}_1^\top I \mathbf{v}_2 \\ &= \mathbf{v}_1 \cdot \mathbf{v}_2 \end{aligned}$$

RM-Q2. Show[1] that the length of a free vector is not changed by rotation, that is, that $||\mathbf{v}|| = ||R\mathbf{v}||$.

Solution

Note $||\mathbf{v}||^2 = \mathbf{v}^\top \mathbf{v} \implies ||\mathbf{v}|| = +\sqrt{\mathbf{v}^\top \mathbf{v}}$. Therefore,

$$\begin{aligned} ||R\mathbf{v}|| = +\sqrt{(R\mathbf{v})^\top (R\mathbf{v})} &= \sqrt{\mathbf{v}^\top R^\top R \mathbf{v}} \\ &= \sqrt{\mathbf{v}^\top \mathbf{v}} = ||\mathbf{v}|| \end{aligned}$$

RM-Q3. Show[1] that the distance between points is not changed by rotation, that is, that $||\mathbf{p}_1 - \mathbf{p}_2|| = ||R\mathbf{p}_1 - R\mathbf{p}_2||$.

[1] Problems selected from Spong, Hutchinson, and Vidyasagar, *Robot Modeling and Control*, Chapter 2.

Solution

This follows from the answer to the previous problem with $\mathbf{v} = \mathbf{p}_1 - \mathbf{p}_2$.

RM-Q4. If[1] a matrix R satisfies $R^\top R = I$, show that the column (and row) vectors or R are of unit length and mutually perpendicular.

Solution

Let $R = [r_1, r_2, r_3]$ where $r_i = \begin{pmatrix} r_{1i} \\ r_{2i} \\ r_{3i} \end{pmatrix}$. Then $R^\top R = I$ implies

$$\begin{bmatrix} r_1^\top r_1 & r_1^\top r_2 & r_1^\top r_3 \\ r_2^\top r_1 & r_2^\top r_2 & r_2^\top r_3 \\ r_3^\top r_1 & r_3^\top r_2 & r_3^\top r_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Equating entries shows that the columns (and rows) of R are of unit length and mutually orthogonal.

RM-Q5. If[1] a matrix R satisfies $R^\top R = I$, then

a) Show that $\det R = \pm 1$. Hint: use properties of determinants.

b) Show that $\det R = +1$ if we restrict ourselves to right-handed coordinate frames. Hint: use the definition of cross product and the fact that cross product takes two vectors and returns a third vector that follows the right hand rule.

Solution

a) For any two matrices A and B, $det(A^\top) = det(A)$ and $det(AB) = det(A)det(B)$. Thus, if R is orthogonal

$$1 = det(I) = det(R^\top R) = det(R^\top)det(R) = det(R)^2$$

which implies that

$$det(R) = \pm\sqrt{1} = \pm 1.$$

b) For a right-handed coordinate system, $r_1 \times r_2 = r_3$. This implies that

$$r_{12}r_{23} - r_{13}r_{22} = r_{31}; \quad -r_{11}r_{23} + r_{13}r_{21} = r_{32}; \;\; r_{11}r_{22} - r_{12}r_{21} = r_{33}.$$

Now, expanding $det(R)$ about column 3 gives:

$$\begin{aligned} det(R) &= det \begin{bmatrix} r_{11} & r_{21} & r_{31} \\ r_{12} & r_{22} & r_{32} \\ r_{13} & r_{23} & r_{33} \end{bmatrix} \\ &= r_{31}(r_{12}r_{23} - r_{13}r_{22}) - r_{32}(r_{11}r_{23} - r_{13}r_{21}) + r_{33}(r_{11}r_{22} - r_{12}r_{21}) \\ &= r_{31}^2 + r_{32}^2 + r_{33}^2 \\ &= ||r_3||^2 \end{aligned}$$

Thus, because $det(R)$ must be positive number, it equals $+1$.

RM-Q6. A[1] **group** is a set X together with an operation $*$ defined on that set such that

- $x_1 * x_2 \in X$ for all $x_1, x_2 \in X$
- $(x_1 * x_2) * x_3 = x_1 * (x_2 * x_3)$
- There exists an element $I \in X$ such that $I * x = x * I = x$ for all $x \in X$.
- For every $x \in X$, there exists some element $y \in X$ such that $x * y = y * x = I$.

Show that $\mathrm{SO}(n)$ with the operation of matrix multiplication is a group.

Solution

First, note that $x \in \mathrm{SO}(n)$ means that $x^\top x = xx^\top = I$ and $det(x) = 1$.

- The first property follows from:

$$(x_1 x_2)^\top (x_1 x_2) = x_2^\top x_1^\top x_1 x_2 = x_2^\top I x_2 + I$$

so

$$x_1 x_2 \in \mathrm{SO}(n) \ \ \forall x_1, x_2 \in \mathrm{SO}(n).$$

- By the associative property of matrix multiplication,

$$(x_1 x_2) x_3 = x_1 (x_2 x_3), \quad \forall x_1, x_2, x_3 \in \mathrm{SO}(n).$$

- The $n \times n$ identity matrix satisfies the third property.
- Since $x^\top x = xx^\top = I$, it follows that $x^\top = x^{-1}$.

RM-Q7. Suppose[1] A is a $2x2$ rotation matrix. In other words $A^\top A = I$ and $\det A = 1$. Show that there exists a unique θ such that A is of the form:

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Solution

Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SO}(2).$$

From Cramer's rule and the fact that $A \in \mathrm{SO}(3)$ we have

$$A^{-1} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

which implies that $a = d$ and $b = -c$. Thus

$$A = \begin{bmatrix} a & -c \\ c & a \end{bmatrix}$$

with $det(A) = 1 = a^2 + c^2$. Define $\theta = tan^{-1}(c/a)$. Then $cos(\theta) = a$ and $sin(\theta) = c$.

RM-Q8. Verify[1] the following equations (where $R_{z,\alpha}$ is a rotation around the z-axis by α degrees):

a) $R_{z,0} = I$

b) $R_{z,\theta} R_{z,\phi} = R_{z,\theta+\phi}$

c) $(R_{z,\theta})^{-1} = R_{z,-\theta}$

Solution

a) Plugging in 0 for θ shows the first equation is true.

b) The second equation can be shown by multiplying out the matrices and substituting trig identities for sums of angles in cosine and sine functions.

c) The third equation follows from the first and second.

RM-Q9. Consider the following sequence of 3D rotations:

1. Rotate by θ around the current z-axis.
2. Rotate by ψ around the current y-axis.
3. Rotate by ϕ around the current x-axis.

Write the matrix product that will give the resulting rotation matrix. (You do not need to perform matrix multiplication).

Solution

$$R = R_{z,\theta} R_{y,\psi} R_{x,\phi}$$

RM-Q10. Consider the following sequence of 3D rotations:

1. Rotate by ϕ around the world x-axis.
2. Rotate by θ around the current z-axis.
3. Rotate by ψ around the world x-axis.

Write the matrix product that will give the resulting rotation matrix. (You do not need to perform matrix multiplication).

Solution

$$R = R_{x,\psi} R_{x,\phi} R_{z,\theta}.$$

Homogeneous Transformations

HT-Q1. Consider[1] the diagram in Figure 2. A robot is set up 1 meter from a table. The table top is 1 meter high and 1 meter square. Coordinate frame 1 is fixed to the edge of the table as shown. A cube measuring 20 cm on one side is placed in the center of the table with coordinate frame 2 established as the center of the cube. A camera is situated directly above the center of block at a height of 2 meters above the table with coordinate frame 3 as shown. Find the homogeneous transformations relating these frames to the base frame of the robot (coordinate frame 0). Find the homogeneous transformation relating frame 2 to the camera frame (frame 3).

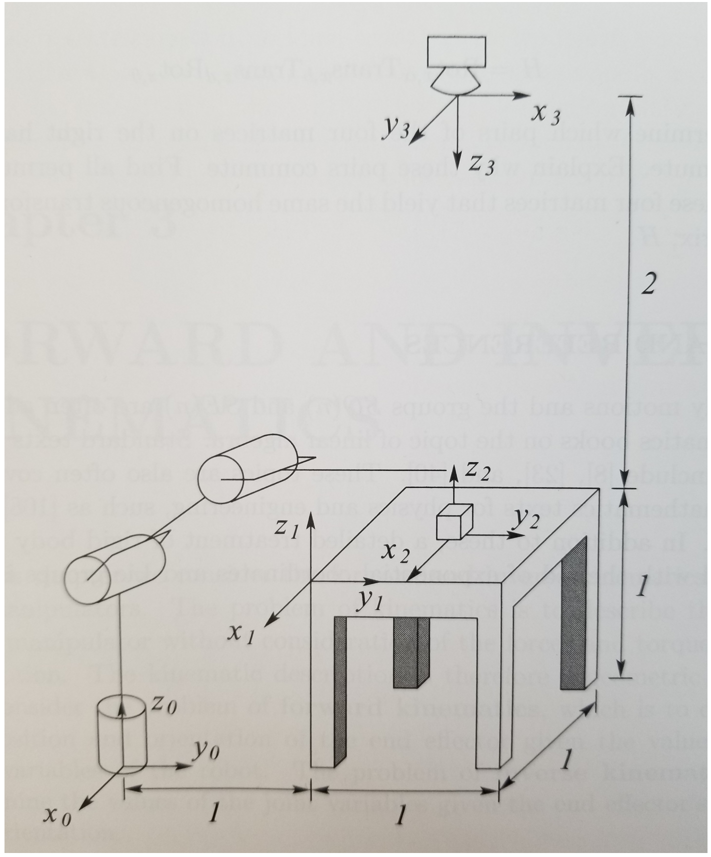


Figure 2: Figure 2-14 from Spong et al. *Robot Modeling and Control*

Solution

$$H_1^0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \; H_2^0 = \begin{bmatrix} 1 & 0 & 0 & -0.5 \\ 0 & 1 & 0 & 1.5 \\ 0 & 0 & 1 & 1.1 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \; H_3^0 = \begin{bmatrix} 0 & 1 & 0 & -0.5 \\ 1 & 0 & 0 & 1.5 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix};$$

$$H_2^3 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1.9 \\ 0 & 0 & 0 & 1 \end{bmatrix};$$

HT-Q2. In[1] the last problem, suppose that, after the camera is calibrated, it is rotated 90^o about z_3. Recompute the above coordinate transformations.

Solution

We first define the new camera frame in terms of the previous one.

$$H^3_{3'} = \begin{bmatrix} cos(\pi/2) & -sin(\pi/2) & 0 & 0 \\ sin(\pi/2) & cos(\pi/2) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

We can then concatenate to find the new transformations:

$$H^0_{3'} = H^0_3 H^3_{3'} = \begin{bmatrix} 0 & 1 & 0 & -0.5 \\ 1 & 0 & 0 & 1.5 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -0.5 \\ 0 & -1 & 0 & 1.5 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$H^{3'}_2 = H^{3'}_3 H^3_2 = (H^3_{3'})^{-1} H^3_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1.9 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1.9 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Note that the inverse of a homogeneous transformation matrix is:

$$\begin{bmatrix} R & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}^{-1} = \begin{bmatrix} R^\top & -R^\top \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$$

HT-Q3. If[1] the block on the table is rotated 90^o about z_2 and moved so that its center has coordinates $[0, 0.8, 0.1]^\top$ (in meters) relative to frame 1, compute the homogeneous transformation that relates the block frame to the camera frame and the transformation that relates the camera frame to the block frame.

Solution

Assuming the original camera location,

$$H^3_{2'} = \begin{bmatrix} 1 & 0 & 0 & 0.3 \\ 0 & -1 & 0 & 0.5 \\ 0 & 0 & -1 & 1.9 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

and

$$H^{2'}_3 = (H^3_{2'})^{-1} = \begin{bmatrix} 1 & 0 & 0 & -0.3 \\ 0 & -1 & 0 & 0.5 \\ 0 & 0 & -1 & 1.9 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$