Rob 501 - Mathematics for Robotics
Recitation #1

Nils Smit-Anseeuw (Courtesy: Wubing Qin)

Sept 12, 2017

1 Notation

Check out this if you want to know more:
`https://gowers.wordpress.com/2011/10/02/basic-logic-relationships-between-statements-negation/`

1. Sets:

   $\mathbb{N}$: Natural numbers, such as $1, 2, 3, \ldots$
   $\mathbb{Z}$: integers, such as $-2, -1, 0, 1, 2, \ldots$
   $\mathbb{Q}$: rational numbers
   $\mathbb{R}$: real numbers
   $\mathbb{C}$: complex numbers
   $\in$: is an element of
   $\notin$: is not an element of
   $\subset$ or $\nsubseteq$: is a subset of
   $\supset$ or $\nsupseteq$: contains
   $\{x : f(x) > 0\}$ or $\{x | f(x) > 0\}$: the set of x or a collection of all the x that satisfies $f(x) > 0$
   $\cup$: the union of two sets
   $\cap$: the intersection of two sets

   Ex: $0 \in \mathbb{Z}$, $\pi \notin \mathbb{Q}$, $\mathbb{Q} \subset \mathbb{R}$, $\mathbb{C} \supset \mathbb{R}$, $2.5 \notin \{x \in \mathbb{R} : x^2 > 10\}$.

2. Logic quantifiers:

   $\forall$: for all, for each, for every, for any
   $\exists$: there exists, there is some, there is at least one
   $\Rightarrow$:implies
   $\Leftrightarrow$: if and only if, iff, is equivalent to
   $\sim$ or $\neg$: negation
   $\vee$: or
   $\wedge$: and

   Ex: $p \implies q$, $p \Leftrightarrow q$, $p \vee q$, $p \wedge q$, $\sim q$, $\sim (p \wedge \sim q)$, $(p \implies q) \Leftrightarrow \big(\sim (p \wedge \sim q)\big)$

3. Others:

   $A \in \mathbb{R}^{m \times n}$: A is an m-by-n matrix of real numbers
   $A \in \mathbb{C}^{m \times n}$: A is an m-by-n matrix of complex numbers
   $[A]_{ij}$: the entry on the i-th row, j-th column of matrix A
   $f : D_1 \to D_2$: a function/mapping/transformation that maps set D_1 to D_2, D_1 is the domain of f and D_2 is the range of f.

Ex:

- A function $f : \mathbb{R} \to \mathbb{R}$ is *continuous at* x_0 if

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x \in \{x : |x - x_0| < \delta\} \Rightarrow |f(x) - f(x_0)| < \epsilon.$$

- A function $f : \mathbb{R} \to \mathbb{R}$ is *continuous on* $\mathbb{R}$ if

$$\forall \epsilon > 0, \forall x \in \mathbb{R}, \exists \delta > 0 : \forall y \in \{z : |z - x| < \delta\} \Rightarrow |f(y) - f(x)| < \epsilon.$$

- A function $f : \mathbb{R} \to \mathbb{R}$ is *uniformly continuous* on $\mathbb{R}$ if

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x, y \in \mathbb{R}, |x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon.$$

2 Matrix

1. Operations:

   (a) Product

      i. Defintion: For matrices $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{m\times l}$, the product $AB \in \mathbb{R}^{n\times l}$ and

$$[AB]_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}, \quad 1 \le i \le n, \quad 1 \le j \le l.$$

      ii. Interpretation
      A matrix pre-multiplied by a row vector is a linear combination of each row of this matrix.

$$\begin{bmatrix} 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 41 & 53 & 62 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 4 \end{bmatrix}, \ \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 4 \end{bmatrix}.$$

      A matrix post-multiplied by a column vector is a linear combination of each column of this matrix.

$$\begin{bmatrix} 0 & 5 & 39 \\ 1 & 0 & 23 \\ 0 & 6 & 11 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 6 \end{bmatrix}, \quad \begin{bmatrix} 0 & 5 \\ 1 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 6 \end{bmatrix}.$$

   (b) Inverse:
   Definition: B is the inverse of A if $AB = BA = I$.

      i. $A^{-1} = \dfrac{\mathrm{adj}(A)}{\det(A)}$

      ii. Elementary row operation $[\ A \mid I\ ] \to [\ I \mid A^{-1}\ ]$

   Ex:

      i. $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1} = \dfrac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$

$$\left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right] \xrightarrow[r_2 - \frac{1}{2} r_1]{\frac{1}{2} r_1} \left[\begin{array}{cc|cc} 1 & 1/2 & 1/2 & 0 \\ 0 & 3/2 & -1/2 & 1 \end{array}\right] \xrightarrow[r_1 - \frac{1}{3} r_2]{\frac{2}{3} r_2} \left[\begin{array}{cc|cc} 1 & 0 & 2/3 & -1/3 \\ 0 & 1 & -1/3 & 2/3 \end{array}\right]$$

ii. $A = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1/5 & 2/5 & 0 \\ -2/5 & -4/5 & 1 \end{bmatrix}, \quad AB = \begin{bmatrix} 1/5 & 2/5 & 0 \\ 2/5 & 4/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; BA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$

iii. $A = \begin{bmatrix} 1 & 2 & 2 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1/5 & -2/5 \\ 2/5 & -4/5 \\ 0 & 1 \end{bmatrix}, \quad AB = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; BA = \begin{bmatrix} 1/5 & 2/5 & 0 \\ 2/5 & 4/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$

Note: A non-square matrix DOES NOT have an inverse.

(c) Transpose: $B = A^\top$ if $b_{ij} = a_{ji}, \forall i, j$.

(d) Properties: For $A, B \in \mathbb{R}^{n \times n}$,

i. $(AB)^\top = B^\top A^\top$

ii. $(A + B)^\top = A^\top + B^\top$

iii. $(AB)^{-1} = B^{-1} A^{-1}$

iv. $(A^\top)^{-1} = (A^{-1})^\top$

2. Determinant

(a) Definition:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij}, \forall i$$

$$M_{ij} : \text{Minor} \qquad\qquad (-1)^{i+j} M_{ij} : \text{cofactor}$$

Ex:

$$\det \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 1 & 3 \end{bmatrix} = 1 \cdot \det \begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix} + 2 \cdot \det \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix} = 1 \cdot 3 + 2 \cdot (-2) = -1.$$

(b) Properties:

i. $\det(kA) = k^n \det(A)$

ii. $\det(A) = \prod_{i=1}^{n} \lambda_i$

iii. $\det(A^\top) = \det(A)$

iv. $\det(A^{-1}) = \dfrac{1}{\det A}$

v. $\det(AB) = \det(A) \det(B)$

3. Trace

(a) Definition:

$$\mathrm{trace}(A) = \sum_{i=1}^{n} a_{ii}$$

(b) Properties

i. $\mathrm{trace}(A) = \sum_{i=1}^{n} \lambda_i$

ii. $\mathrm{trace}(A + B) = \mathrm{trace}(A) + \mathrm{trace}(B)$

iii. $\mathrm{trace}(AB) = \mathrm{trace}(BA)$

iv. $\mathrm{trace}(ABC) = \mathrm{trace}(BCA) = \mathrm{trace}(CAB)$

Ex: $A = \begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

$$\begin{aligned}
A + B &= \begin{bmatrix} 3 & 5 \\ 4 & 2 \end{bmatrix} \\
\text{trace}(A) &= 2 + 1 = 3 \\
\text{trace}(B) &= 1 + 1 = 2 \\
\text{trace}(A + B) &= 3 + 2 = 5 \\
AB &= \begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 7 \\ 4 & 5 \end{bmatrix} \\
BA &= \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 6 & 7 \end{bmatrix} \\
\text{trace}(AB) &= 8 + 5 = 13 \\
\text{trace}(BA) &= 6 + 7 = 13
\end{aligned}$$

3 Minimization

1. Without constraint: Derivatives/Gradient

   Ex: $f(x, y) = x^2 + y^2 + 3xy + x - y$.

   Soln:

   $$\text{Necessary conditions} \Rightarrow \begin{cases} \frac{\partial}{\partial x} f(x, y) & = 2x + 3y + 1 = 0 \\ \frac{\partial}{\partial y} f(x, y) & = 2y + 3x - 1 = 0 \end{cases} \Rightarrow \begin{cases} x = 1 \\ y = -1 \end{cases}$$

   $\nabla^2 f(x, y) = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix} \not\succeq 0 \Rightarrow (1, -1)$ is not minimal. Since this is the only point satisfying the necessary conditions for optimality, we know this function has no minimizer. In fact, this function is unbounded from below.

2. With equality constraints: Lagrange Multipliers

   Ex: $f(x, y) = 2x - y$ s.t. $x^2 + \frac{1}{4} y^2 = 2$.

   Soln: Let $g(x,\, y,\, \lambda) = 2x - y + \lambda\, (x^2 + \frac{1}{4} y^2 - 2)$.

   $$\text{Necessary conditions} \Rightarrow \begin{cases} \frac{\partial}{\partial x} g(x,\, y,\, \lambda) & = 2 + 2\lambda x = 0 \\ \frac{\partial}{\partial y} g(x,\, y,\, \lambda) & = -1 + \frac{1}{2} \lambda y = 0 \\ \frac{\partial}{\partial \lambda} g(x,\, y,\, \lambda) & = x^2 + \frac{1}{4} y^2 - 2 = 0 \end{cases} \Rightarrow \begin{cases} x = 1 \\ y = -2 \\ \lambda = -1 \end{cases} \text{or} \begin{cases} x = -1 \\ y = 2 \\ \lambda = 1 \end{cases}$$

4 Probability

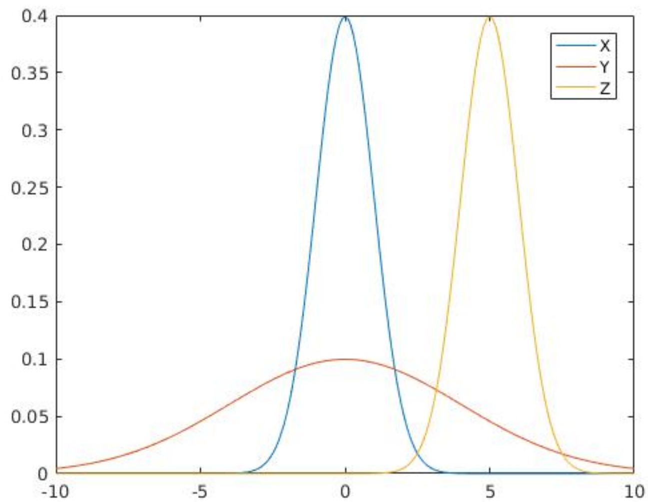
1. Basics

   Probability density function (PDF):
   $$f_X(x) := \frac{\mathrm{d}}{\mathrm{d}x} F_X(x)$$
   Cumulative distribution function (CDF):
   $$F_X(x) := P\big(X \leq x\big)$$
   Conditional probability:
   $$P(A|B) = \frac{P(A \cap B)}{P(B)}$$
   Theorem of total probability:

   Suppose $\{A_1, \dots, A_n\}$ is a partition of the sample space Ω, i.e., $\forall i, j : A_i \cap A_j = \emptyset$, $\bigcup_{i=1}^{n} A_i = \Omega$, then the probability of event B is
   $$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$
   Bayes rule:
   $$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$
   Marginal PDF:
   $$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad f_{X_1}(x_1) = \int_{\mathbb{R}} f_X(x)\mathrm{d}x_2$$
   Conditional pdf:
   $$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad f_{X_1|X_2}(x_1|X_2 = a) = \frac{f_{X_1X_2}(x_1, a)}{f_{X_2}(a)}$$

2. normal distribution

   - Scalar case: $X \sim N(\mu, \sigma^2)$
   $$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\big\{ -\frac{(x-\mu)^2}{2\sigma^2} \big\}$$
   - Vector case: $X \sim N(\mu, \Sigma)$
   $$f_X(x) = \frac{1}{\sqrt{\det(2\pi\Sigma)}} \exp\big\{ -\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu) \big\}$$
   - **Standard Normal Distribution**
     Special case of normal distribution with $\mu = 0$ and $\sigma = 1$
     Ex: Figure shows a standard normal distribution, $X \sim N(0, 1)$. Overlay the plot for $Y = 4X$ and $Z = X + 5$


0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-10
-5
0
5
10
X
Y
Z