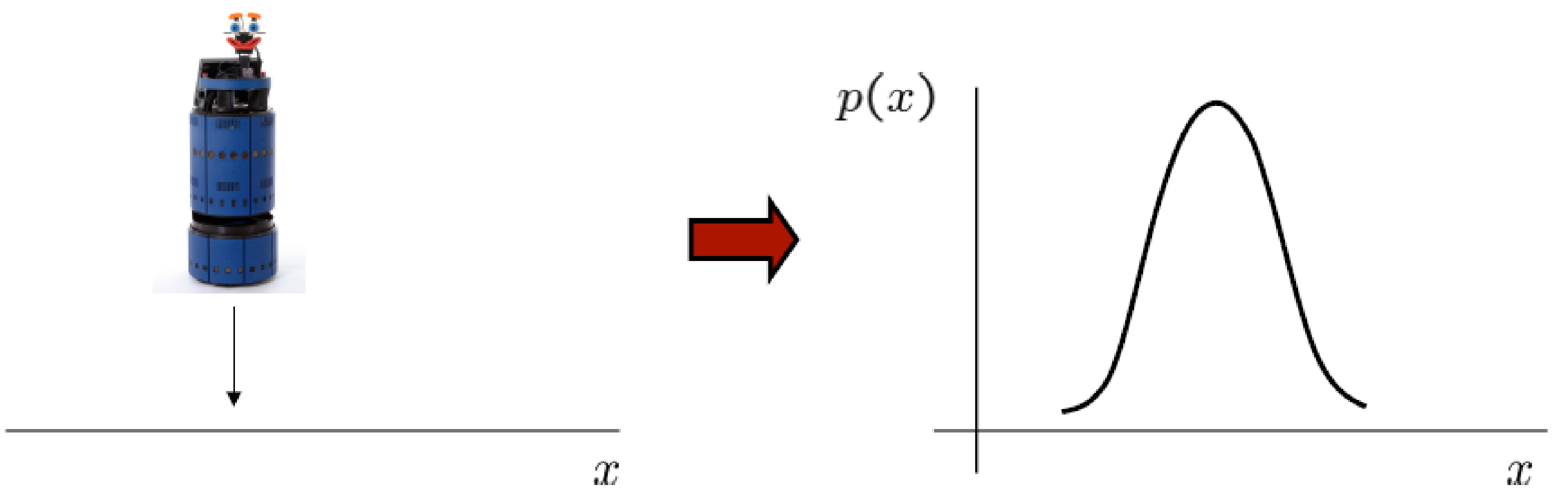
Probability Primer

Nived Chebrolu

Slides adapted from Cyrill Stachniss at University of Bonn.

Probabilistic Approaches

- Uncertainty in robot motion and observations
- Use of probability theory to explicitly represent the uncertainty



"The robot is exactly here"

"The robot is somewhere here"

Probability Primer

Why Probabilities?

- Probability theory is the key tool for robust state estimation and robotics
- Probability theory allows to explicitly model uncertainties
- Roughly 90% of the techniques presented in this course rely on it

Get comfortable with probability theory

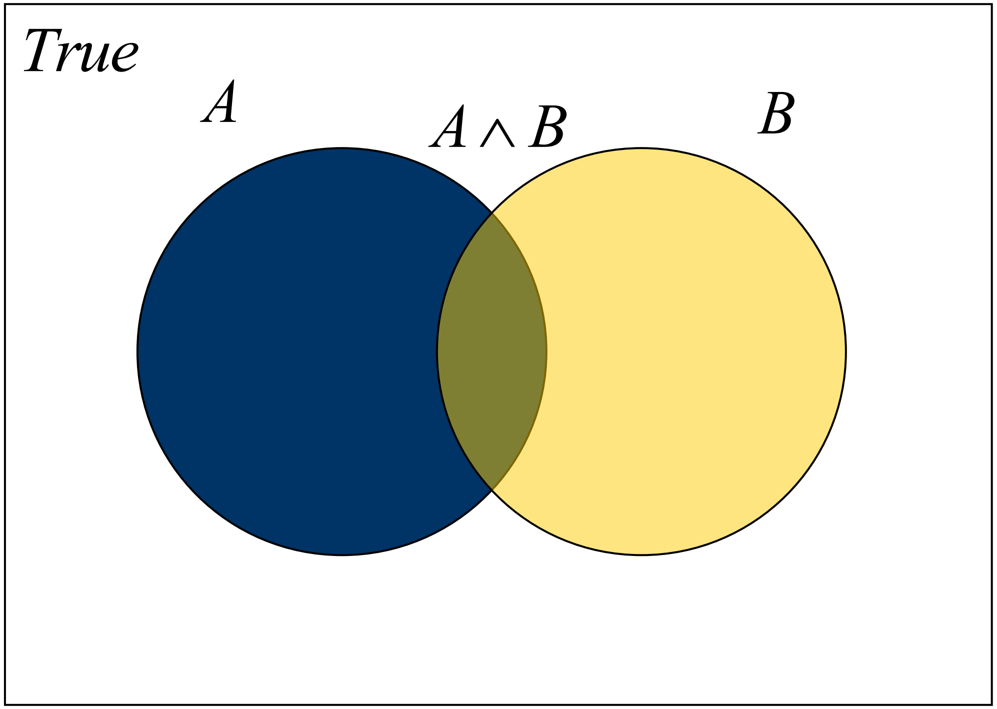
Axioms of Probability Theory

Pr*(A)* denotes probability that proposition *A* is true.

- $0 \leq \Pr(A) \leq 1$
- $\Pr(\mathit{True}) = 1$
- $\Pr(\mathit{False}) = 0$
- $\Pr(A \vee B) = \Pr(A) + \Pr(B) - \Pr(A \wedge B)$

A Closer Look at Axiom 3

$$\Pr(A \vee B) = \Pr(A) + \Pr(B) - \Pr(A \wedge B)$$



Example for Using the Axioms

Based on the axioms, we can derive the probability of *not A:*

$$\Pr(\neg A)$$

Example for Using the Axioms

Based on the axioms, we can derive the probability of *not A:*

$$\Pr(A \vee \neg A) = \Pr(A) + \Pr(\neg A) - \Pr(A \wedge \neg A)$$

Example for Using the Axioms

Based on the axioms, we can derive the probability of *not A:*

$$\begin{aligned}\Pr(A \vee \neg A) &= \Pr(A) + \Pr(\neg A) - \Pr(A \wedge \neg A) \\ \Pr(\mathit{True}) &= \Pr(A) + \Pr(\neg A) - \Pr(\mathit{False})\end{aligned}$$

Example for Using the Axioms

Based on the axioms, we can derive the probability of *not A:*

$$
\begin{aligned}
\Pr(A \vee \neg A) &= \Pr(A) + \Pr(\neg A) - \Pr(A \wedge \neg A) \\
\Pr(\mathit{True}) &= \Pr(A) + \Pr(\neg A) - \Pr(\mathit{False}) \\
1 &= \Pr(A) + \Pr(\neg A) - 0
\end{aligned}
$$

Example for Using the Axioms

Based on the axioms, we can derive the probability of *not A:*

$$\begin{aligned}
\Pr(A \vee \neg A) &= \Pr(A) + \Pr(\neg A) - \Pr(A \wedge \neg A) \\
\Pr(\mathit{True}) &= \Pr(A) + \Pr(\neg A) - \Pr(\mathit{False}) \\
1 &= \Pr(A) + \Pr(\neg A) - 0 \\
\Pr(\neg A) &= 1 - \Pr(A)
\end{aligned}$$

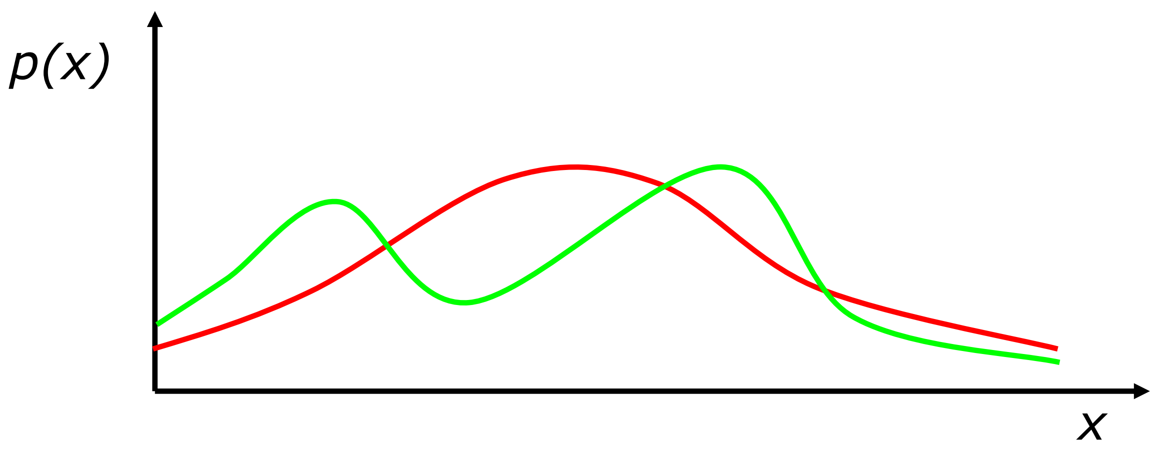
Discrete Random Variables

- X denotes a **random variable**
- X can take on a countable number of values in $\{x_1, x_2, \ldots, x_n\}$
- $P(X = x_i) = P(x_i)$ is the **probability** that the random variable X takes the value x_i
- $P()$ is called **probability mass function**

Continuous Random Variables

- X takes on values in the continuum
- $p(X = x) = p(x)$ is a **probability density function**

$$Pr(x \in [a, b]) = \int_a^b p(x)\, dx$$



Probability Sums up to One

Discrete case

$$\sum_{x_i} P(x_i) = 1$$

Continuous case

$$\int p(x)\, dx = 1$$

Joint and Conditional Probability

- $P(X = x \text{ and } Y = y) = P(x, y)$
- If X and Y are **independent** then

$$P(x, y) = P(x)\,P(y)$$

- $P(x \mid y)$ is the probability of ***X* given *Y***

$$P(x, y) = P(x \mid y)\,P(y)$$

$$P(x \mid y) = P(x, y)/P(y)$$

- If X and Y are **independent** then

$$P(x \mid y) = P(x)$$

Law of Total Probability

Discrete case

$$P(x) = \sum_y P(x \mid y)\, P(y)$$

Continuous case

$$p(x) = \int p(x \mid y)\, p(y)\, dy$$

Marginalization

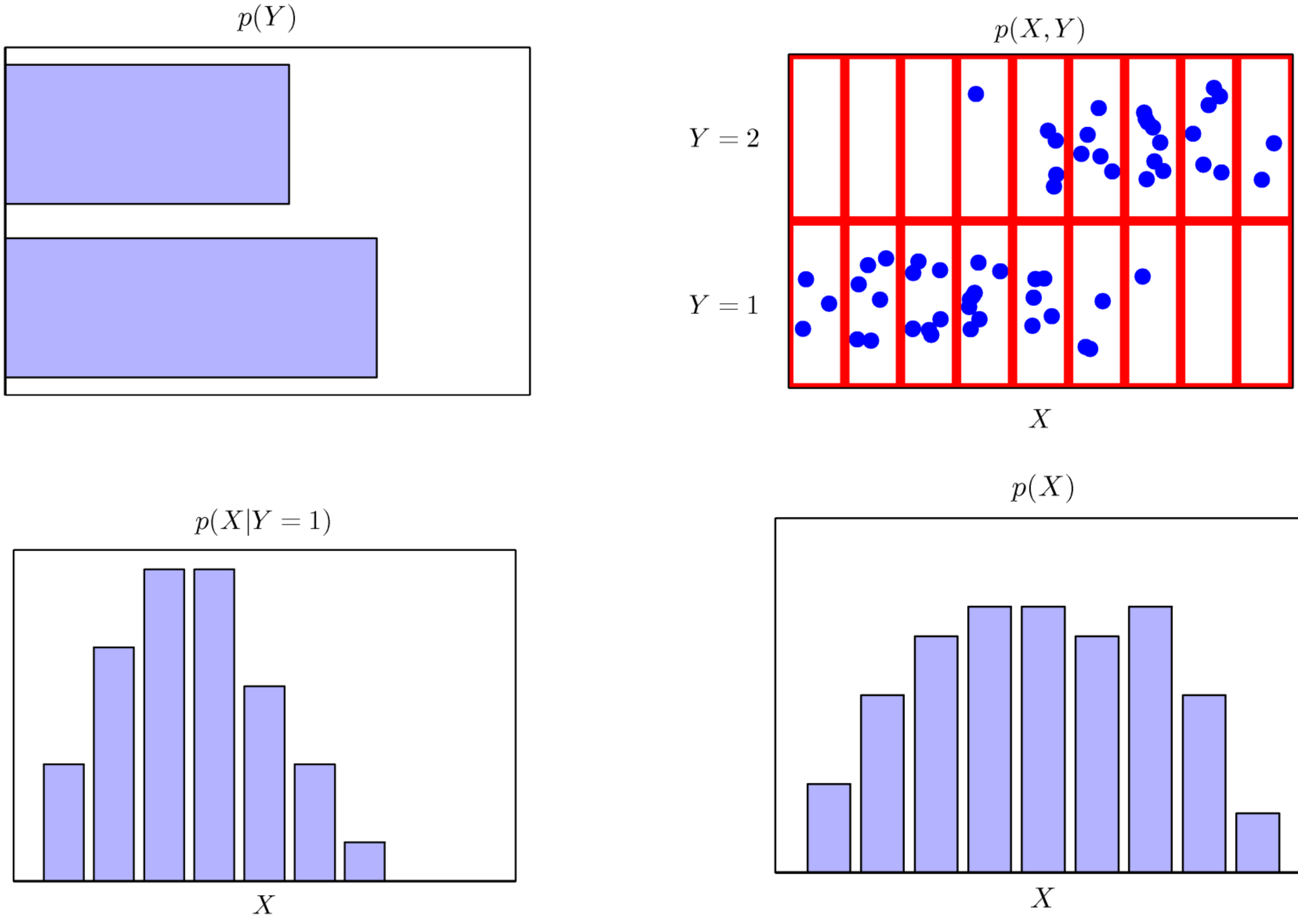
Discrete case

$$P(x) = \sum_y P(x, y)$$

Continuous case

$$p(x) = \int p(x, y)\, dy$$

Example



Example

	M	$\neg M$	$\sum$
S	20	20	
$\neg S$	35	45	

Example

	M	$\neg M$	$\sum$
S	20	20	40
$\neg S$	35	45	80
$\sum$			

Example

	M	$\neg M$	$\sum$
S	20	20	40
$\neg S$	35	45	80
$\sum$	55	65	120

Example

	M	$\neg M$	$\sum$
S	20	20	40
$\neg S$	35	45	80
$\sum$	55	65	120

$$P(M, S) =$$

Example

	M	$\neg M$	$\sum$
S	20	20	40
$\neg S$	35	45	80
$\sum$	55	65	120

$$P(M, S) = \frac{20}{120} \qquad P(M, \neg S) =$$

Example

	M	$\neg M$	$\sum$
S	20	20	40
$\neg S$	35	45	80
$\sum$	55	65	120

$$P(M,S) = \frac{20}{120} \qquad P(M,\neg S) = \frac{35}{120} \qquad \ldots$$

Example

	M	$\neg M$	$\sum$
S	20	20	40
$\neg S$	35	45	80
$\sum$	55	65	120

$$P(M, S) = \frac{20}{120} \qquad P(M, \neg S) = \frac{35}{120} \qquad \ldots$$

$$P(M) = \qquad P(\neg M) = \qquad P(S) =$$

Example

	M	$\neg M$	$\sum$
S	20	20	40
$\neg S$	35	45	80
$\sum$	55	65	120

$$P(M, S) = \frac{20}{120} \qquad P(M, \neg S) = \frac{35}{120} \qquad \ldots$$

$$P(M) = \frac{55}{120} \qquad P(\neg M) = \frac{65}{120} \qquad P(S) = \frac{40}{120} \qquad \ldots$$

Example

	M	$\neg M$	$\sum$
S	20	20	40
$\neg S$	35	45	80
$\sum$	55	65	120

$$P(M,S) = \frac{20}{120} \qquad P(M,\neg S) = \frac{35}{120} \qquad \ldots$$

$$P(M) = \frac{55}{120} \qquad P(\neg M) = \frac{65}{120} \qquad P(S) = \frac{40}{120} \qquad \ldots$$

$$P(M \mid S) =$$

Example

	M	$\neg M$	$\sum$
S	20	20	40
$\neg S$	35	45	80
$\sum$	55	65	120

$$P(M, S) = \frac{20}{120} \qquad P(M, \neg S) = \frac{35}{120} \qquad \ldots$$

$$P(M) = \frac{55}{120} \qquad P(\neg M) = \frac{65}{120} \qquad P(S) = \frac{40}{120} \qquad \ldots$$

$$P(M \mid S) = \frac{P(M, S)}{P(S)} =$$

Example

	M	$\neg M$	$\sum$
S	20	20	40
$\neg S$	35	45	80
$\sum$	55	65	120

$$P(M, S) = \frac{20}{120} \qquad P(M, \neg S) = \frac{35}{120} \qquad \ldots$$

$$P(M) = \frac{55}{120} \qquad P(\neg M) = \frac{65}{120} \qquad P(S) = \frac{40}{120} \qquad \ldots$$

$$P(M \mid S) = \frac{P(M, S)}{P(S)} = \frac{20/120}{40/120} = \frac{20}{40} \qquad \ldots$$

Bayes' Rule

$$P(x,y) = P(x \mid y)\,P(y)$$

$$P(x,y) = P(y \mid x)P(x)$$

$$\Downarrow$$

$$\boxed{P(x \mid y) = \frac{P(y \mid x)\,P(x)}{P(y)} = \frac{\text{likelihood}\,\text{prior}}{\text{evidence}}}$$

Bayes Rule
with Background Knowledge

$$P(x \mid y) = \frac{P(y \mid x)\, P(x)}{P(y)}$$

$$P(x \mid y, z) = \frac{P(y \mid x, z)\, P(x \mid z)}{P(y \mid z)}$$

Conditional Independence

Two variables are conditional independent if

$$P(x, y \mid z) = P(x \mid z)\, P(y \mid z)$$

- Equivalent to $P(x \mid z) = P(x \mid z, y)$
  and $P(y \mid z) = P(y \mid z, x)$

- But this does **not imply**

$$P(x, y) = P(x)\, P(y)$$

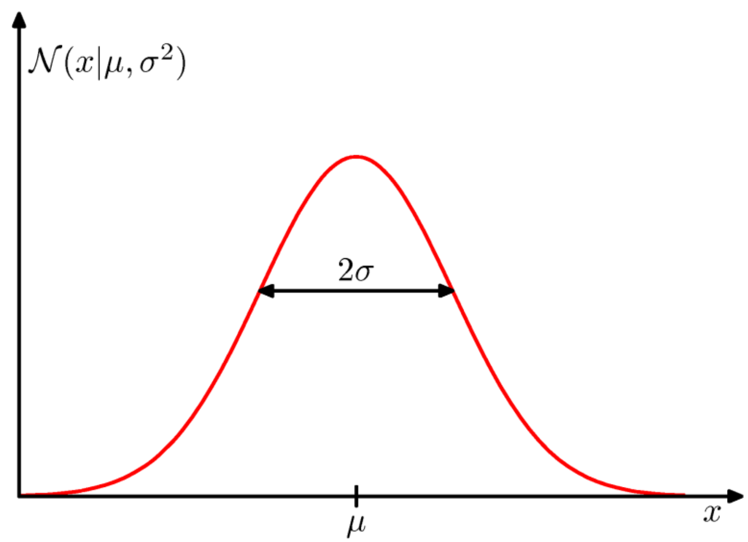
(independence/marginal independence)

Normal Distribution

- Univariate: 1-dimensional
- Unimodal: one mode

$$\mathcal{N}\left(x|\mu,\sigma^2\right) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

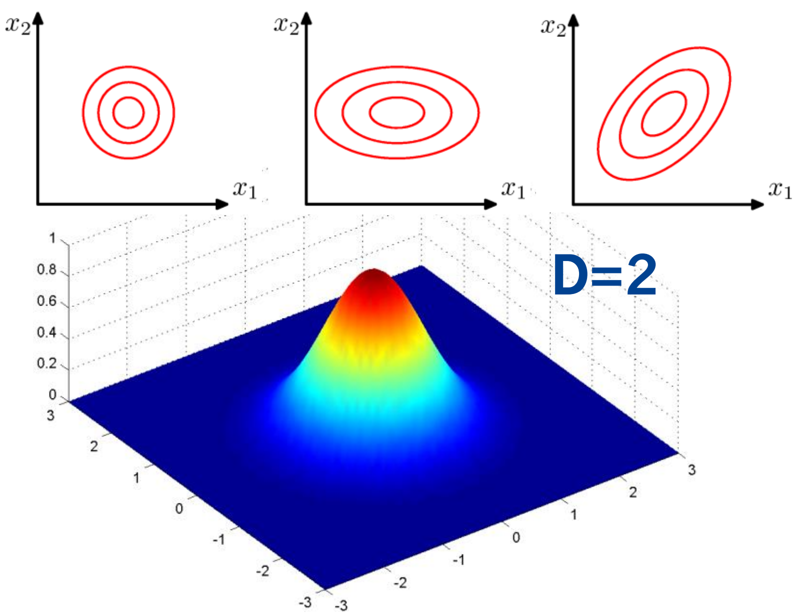
37

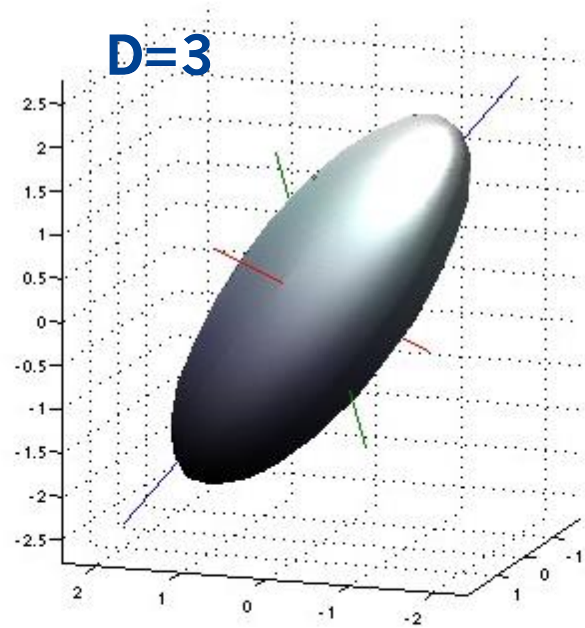


Multivariate Normal Distribution

- Multivariate: multi-dimensional
- Unimodal: one mode

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}$$

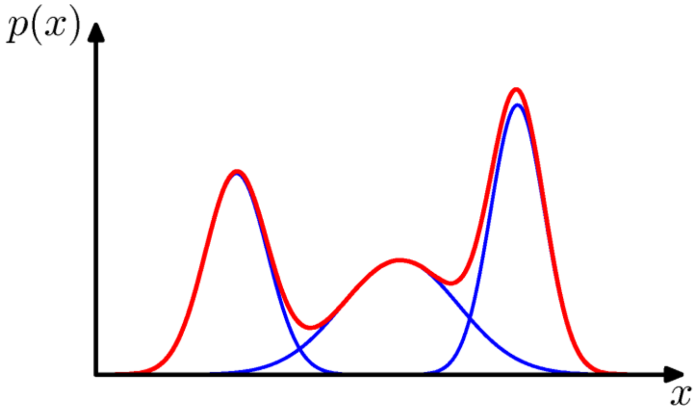


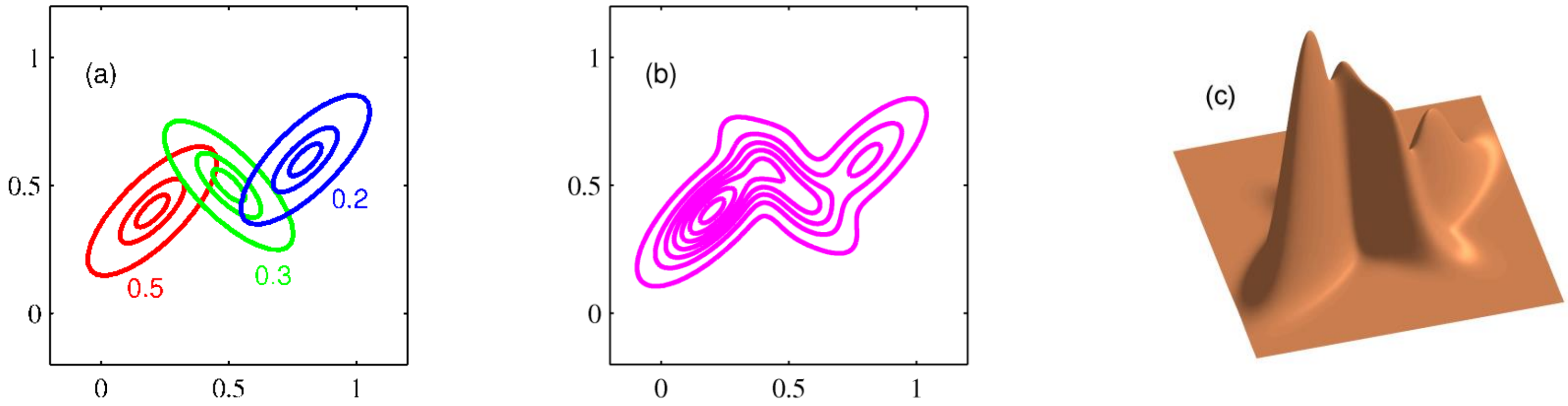


Gaussian Mixture

- Multi-modal (number of components)

$$p(\mathbf{x}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$$





Summary

- Probability is an essential tool to solve state estimation problems
- Make sure you know how to deal with
  - probabilities,
  - uncertainties,
  - Bayes' rule,
  - …