Lecture #2

Basic probability

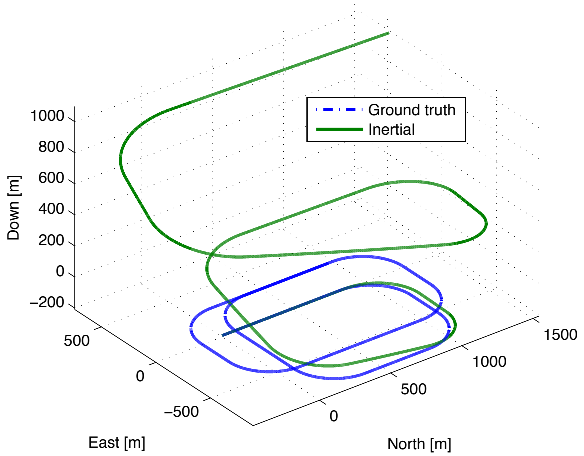
Bayesian inference

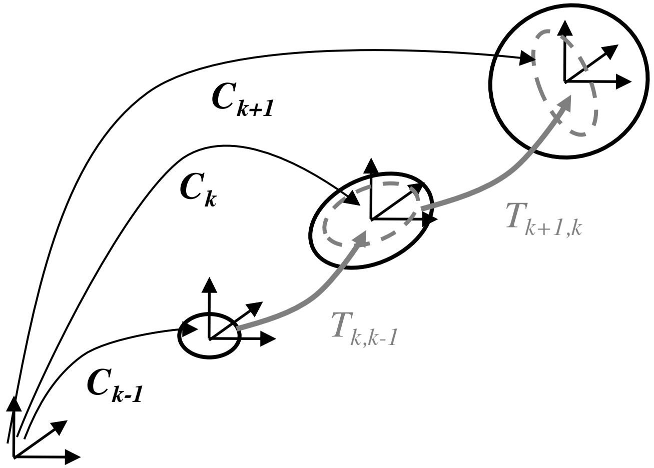
Extended Kalman filter (EKF)

Previous Lecture

- Dead reckoning: e.g. inertial navigation, visual odometry
- Drift accumulates over time (frame-to-frame motion estimation)
- How to improve estimation accuracy?
  - Incorporate additional sensors
  - Use prior information if exists (how accurate?)
  - Estimate motion using also past states

Involves probabilistic inference





Objectives of this Lecture

- Refresh basic probability
- Understand marginalization and conditioning concepts

- Learn key concepts in
  - Bayesian inference
  - (Extended) Kalman filter and smoothing

Course Diagram

3D Transformations | Dead Reckoning | Basic Probability

Bayesian Inference | Extended Kalman/Information Filter

Projective camera geometry | Multi View Geometry | Feature Matching

Bundle Adjustment | VAN, SLAM | Graphical models

Incremental Smoothing and Mapping (iSAM)

Advanced topics (subject to progress in class)

Multi-Robot SLAM & VAN | Belief Space Planning

Discrete Random Variables

- X denotes a random variable
- x denotes a specific value that X may assume
- $P(X=x)$ is the probability the random variable X has value x.
- Discrete probabilities sum to 1: $\sum_x p(X=x) = 1$

Example: coin flip

- X can take on the values <u>heads</u> or <u>tails</u>
- A fair coin is characterized by: $p(X=head) = p(X=tail) = 0.5$

Continuous Random Variables

- X can take on continuous values
- $p(X=x)$ is called the **probability density function** (pdf)
  - To simplify notations, abbreviate: $p(x)$

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

- A pdf always sums to 1:

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

Joint and Conditional Probabilities

- Joint distribution of two random variables X and Y

$$p\left(x,y\right)=p\left(X=x \text{ and } Y=y\right)$$

- If X and Y are independent:

$$p\left(x,y\right)=p\left(x\right)p\left(y\right)$$

- Conditional probability – probability of x given y:

$$p\left(x|y\right)=\frac{p\left(x,y\right)}{p\left(y\right)}$$

- If X and Y are independent:

$$p\left(x|y\right)=p\left(x\right)$$

Conditional Independence

- Definition: $p\left(x, y|z\right) = p\left(x|z\right) p\left(y|z\right)$

- The same as

$$p\left(x|z\right) = p\left(x|y, z\right) \;\; , \;\; p\left(y|z\right) = p\left(y|x, z\right)$$

- Does <u>not</u> necessarily mean X and Y are independent

Marginalization

- Also known as the Law of Total Probability:
  - Discrete case:
    $$p(x) = \sum_y p(x, y)$$
    $$= \sum_y p(x|y)\, p(y)$$
  - Continuous case:
    $$p(x) = \int p(x, y)\, dy$$
    $$= \int p(x|y)\, p(y)\, dy$$

Marginalization - Example

- Discrete case, 2 random variables, each can assume 4 values

$p(x, y)$:

	x_1	x_2	x_3	x_4
y_1	$4/32$	$2/32$	$1/32$	$1/32$
y_2	$2/32$	$4/32$	$1/32$	$1/32$
y_3	$2/32$	$2/32$	$2/32$	$2/32$
y_4	$8/32$	0	0	0

Recall

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

$$p(y) = \sum_x p(x, y)$$

$$\sum_x \sum_y p(x, y) = 1$$

Marginalization - Example

- Discrete case, 2 random variables, each can assume 4 values

$p(x, y)$:

	x_1	x_2	x_3	x_4
y_1	$4/32$	$2/32$	$1/32$	$1/32$
y_2	$2/32$	$4/32$	$1/32$	$1/32$
y_3	$2/32$	$2/32$	$2/32$	$2/32$
y_4	$8/32$	0	0	0
$p(x)$	$16/32$	$8/32$	$4/32$	$4/32$

Recall

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

$$p(y) = \sum_x p(x, y)$$

$$\sum_x \sum_y p(x, y) = 1$$

Marginalization - Example

- Discrete case, 2 random variables, each can assume 4 values

$p(x,y):$

	x_1	x_2	x_3	x_4	$p(y)$
y_1	$4/32$	$2/32$	$1/32$	$1/32$	$8/32$
y_2	$2/32$	$4/32$	$1/32$	$1/32$	$8/32$
y_3	$2/32$	$2/32$	$2/32$	$2/32$	$8/32$
y_4	$8/32$	0	0	0	$8/32$
$p(x)$	$16/32$	$8/32$	$4/32$	$4/32$	$32/32$

Recall

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

$$p(y) = \sum_x p(x,y)$$

$$\sum_x \sum_y p(x,y) = 1$$

Expectation of a Random Variable

- Discrete case: $$\mathbb{E}\left[X\right] = \sum_i x_i p\left(x_i\right)$$
- Continuous case: $$\mathbb{E}\left[X\right] = \int x p\left(x\right) dx$$
- The expected value is the weighted average of all values a random variable can take on
- Expectation is a linear operator:

$$\mathbb{E}\left[aX + b\right] = a\mathbb{E}\left[X\right] + b$$

Covariance of a Random Variable

- The covariance measures the squared expected deviation from the mean (prove!):

  – Scalar case: $Cov\left[X\right] = \mathbb{E}\left[X - \mathbb{E}\left[X\right]\right]^2 = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2$

  – Multivariate case: $X \in \mathbb{R}^n$

$$Cov\left[X\right] = \mathbb{E}\left[\left(X - \mathbb{E}\left[X\right]\right)\left(X - \mathbb{E}\left[X\right]\right)^T\right]$$

Bayes and Chain Rules

- Bayes rule (more soon):

$$\boxed{p\left(x|y\right) = \frac{p\left(y|x\right)p\left(x\right)}{p\left(y\right)}}$$

discrete

$$= \frac{p\left(y|x\right)p\left(x\right)}{\sum_{x'} p\left(y|x'\right)p\left(x'\right)}$$

continuous

$$= \frac{p\left(y|x\right)p\left(x\right)}{\int p\left(y|x'\right)p\left(x'\right)dx'}$$

- Chain rule (already appeared):

$$p\left(x,y\right) = p\left(x|y\right)p\left(y\right)$$

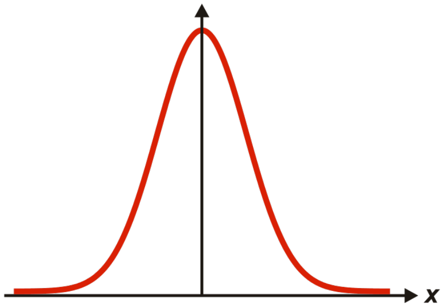
Gaussian Distributions

Gaussian Distribution

- The pdf of a one-dimensional Gaussian distribution with mean μ and variance σ^2:

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

- Notation: $N(x;\mu,\sigma^2)$ or $x \sim N(\mu,\sigma^2)$



Gaussian Distribution

- Multivariate Gaussian distribution (x is a vector)

$$p\left(x\right)=\frac{1}{\sqrt{\det\left(2\pi\Sigma\right)}}\exp\left(-\frac{1}{2}\left(x-\mu\right)^{T}\Sigma^{-1}\left(x-\mu\right)\right)$$

  - μ : mean vector
  - Σ : covariance matrix (positive semidefinite, symmetric)

  - Squared Mahalanobis norm: $\left\|x-\mu\right\|_{\Sigma}^{2}\doteq\left(x-\mu\right)^{T}\Sigma^{-1}\left(x-\mu\right)$

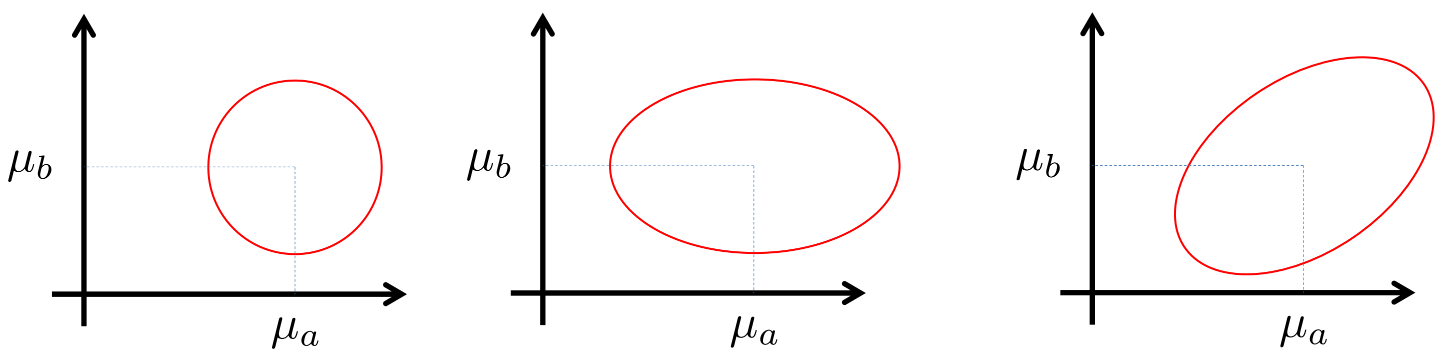
$$\Rightarrow\quad p\left(x\right)=\frac{1}{\sqrt{\det\left(2\pi\Sigma\right)}}\exp\left(-\frac{1}{2}\left\|x-\mu\right\|_{\Sigma}^{2}\right)=N\left(\mu,\Sigma\right)$$

Example – Covariance, 2D Case

- Two-dimensional state $x \doteq \begin{pmatrix} a \\ b \end{pmatrix}$
- Joint Gaussian pdf: $p(x) = N(\mu, \Sigma)$

$$\mu \doteq \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix} \qquad \Sigma \doteq \begin{bmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ab}^T & \Sigma_{bb} \end{bmatrix}$$

- Intuition:



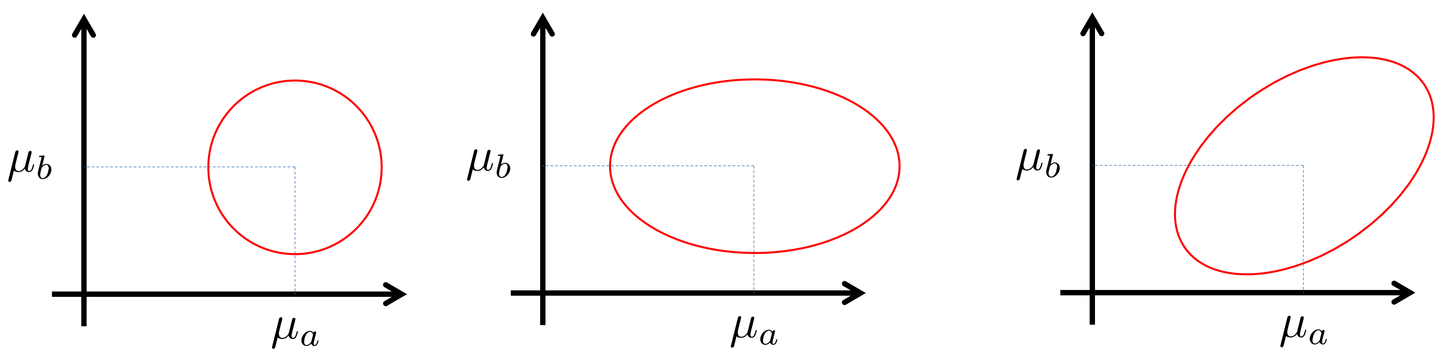
$\Sigma_{aa}, \Sigma_{ab}, \Sigma_{bb} = ?$ $\qquad$ $\Sigma_{aa}, \Sigma_{ab}, \Sigma_{bb} = ?$ $\qquad$ $\Sigma_{aa}, \Sigma_{ab}, \Sigma_{bb} = ?$

Example – Covariance, 2D Case

- Two-dimensional state $x \doteq \begin{pmatrix} a \\ b \end{pmatrix}$
- Joint Gaussian pdf: $p(x) = N(\mu, \Sigma)$

$$\mu \doteq \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix} \qquad \Sigma \doteq \begin{bmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ab}^T & \Sigma_{bb} \end{bmatrix}$$

- Intuition:



$\Sigma_{aa} = \Sigma_{bb}$

$\Sigma_{aa} > \Sigma_{bb}$
$\Sigma_{ab} = 0$

$\Sigma_{ab} \neq 0$

Gaussian Distribution – Information Form

- Covariance form (previous slide): $x \sim N\left(x; \mu, \Sigma\right)$
- Information form: $x \sim N^{-1}\left(x; \eta, \Lambda\right)$
  - Information matrix: $\Lambda \doteq \Sigma^{-1}$
  - Information vector: $\eta \doteq \Lambda \mu$

- It can be shown that:

$$p\left(x\right) = N^{-1}\left(\eta, \Lambda\right) = \frac{\exp\left(-\frac{1}{2}\eta^T \Lambda^{-1} \eta\right)}{\sqrt{\det\left(2\pi\Lambda^{-1}\right)}} \exp\left(-\frac{1}{2} x^T \Lambda x + \eta^T x\right)$$

Back to Conditioning and Marginalization

- Covariance form: marginalization is easy
- Information form: conditioning is easy

- Recall:
  - Marginalization $$p\left(x\right)=\int p\left(x,y\right)dy$$
  - Conditioning $$p\left(x|y\right)=\frac{p\left(x,y\right)}{p\left(y\right)}$$

Back to Conditioning and Marginalization

$$p\left(x,y\right)=\mathcal{N}\left(\left[\begin{array}{c}\mu_x\\ \mu_y\end{array}\right],\left[\begin{array}{cc}\Sigma_{xx} & \Sigma_{xy}\\ \Sigma_{yx} & \Sigma_{yy}\end{array}\right]\right)=\mathcal{N}^{-1}\left(\left[\begin{array}{c}\eta_x\\ \eta_y\end{array}\right],\left[\begin{array}{cc}I_{xx} & I_{xy}\\ I_{yx} & I_{yy}\end{array}\right]\right)$$

Back to Conditioning and Marginalization

$$p(x,y) = \mathcal{N}\left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}\right) = \mathcal{N}^{-1}\left(\begin{bmatrix} \eta_x \\ \eta_y \end{bmatrix}, \begin{bmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{bmatrix}\right)$$

	Marginalization	**Conditioning**
	$p(x) = \int p(x,y)\, dy \doteq \mathcal{N}(\mu, \Sigma) \doteq \mathcal{N}^{-1}(\eta, I)$	$p(x\|y) = \frac{p(x,y)}{p(y)} \doteq \mathcal{N}(\mu', \Sigma') \doteq \mathcal{N}^{-1}(\eta', I')$
Covariance form	$\mu = \mu_x$ $\Sigma = \Sigma_{xx}$	$\mu' = \mu_x + \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y)$ $\Sigma' = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}$
Information form	$\eta = \eta_x - I_{xy}I_{yy}^{-1}\eta_y$ $I = I_{xx} - I_{xy}I_{yy}^{-1}I_{yx}$	$\eta' = \eta_x - I_{xy}y$ $I' = I_{xx}$

Table inspired by Eustice06tro

Recap – Summary Thus Far

- Discrete and continuous random variables
- Joint and conditional probabilities
- Marginalization
- Chain rule
- Bayes rule
- Gaussian distribution

Course Diagram

3D Transformations | Dead Reckoning | Basic Probability

Bayesian Inference | Extended Kalman/Information Filter

Projective camera geometry | Multi View Geometry | Feature Matching

Bundle Adjustment | VAN, SLAM | Graphical models

Incremental Smoothing and Mapping (iSAM)

Advanced topics (subject to progress in class)

Multi-Robot SLAM & VAN | Belief Space Planning

Probabilistic Inference

- Probabilistic inference (or state estimation):
  - Objective: estimate the state x
    - Platform/camera pose
    - Environment model (what do we see?)
    - etc.
  - Given:
    - Sensor measurements z
    - Controls (or odometry readings)
- Bayes rule plays a key role

Probabilistic Inference

- $p(x)$ - Prior probability: Represents knowledge regarding x before incorporating sensor reading(s) z

- $p(x|z)$ - Posterior (or a posteriori) probability: Represents knowledge after incorporating sensor reading(s) z

- **Bayes rule** allows to express $p(x|z)$ via $p(z|x)$ and $p(x)$

Measurement likelihood

Prior

$$p\left(x|z\right) = \frac{p\left(z|x\right)p\left(x\right)}{p\left(z\right)} = \eta p\left(z|x\right)p\left(x\right) \propto p\left(z|x\right)p\left(x\right)$$

Prior on sensor observations, does not depend on x

normalizer

Example

- Quadcopter seeks a landing zone
- Landing zone is marked with many bright lamps
- Quadcopter has a brightness sensor

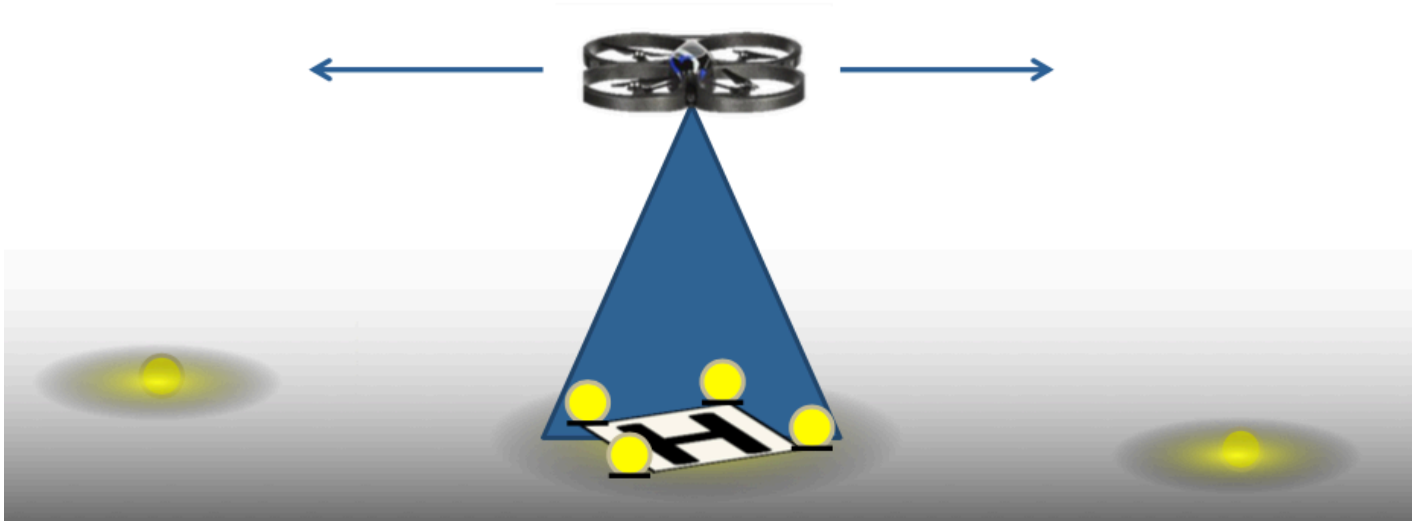


Figure by Dr. J. Sturm

Example

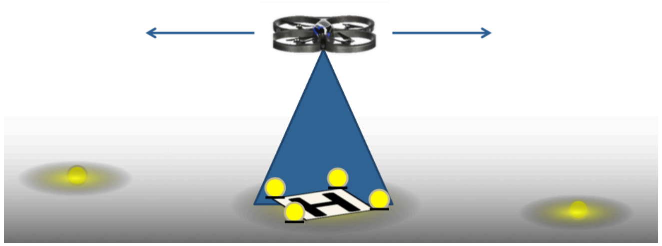
- Binary sensor: $z \in \{\text{bright}, \neg\text{bright}\}$
- Binary state: $x \in \{\text{home}, \neg\text{home}\}$
- Prior on state: $p\left(x = \text{home}\right) = 0.5$

- Sensor model: $p\left(z = \text{bright} \mid x = \text{home}\right) = 0.6$ $\quad p\left(z = \neg\text{bright} \mid x = \text{home}\right) = 0.4$

  $p\left(z = \text{bright} \mid x = \neg\text{home}\right) = 0.3$ $\quad p\left(z = \neg\text{bright} \mid x = \neg\text{home}\right) = 0.7$

- Assume robot observes light: $z = \text{bright}$
- What is the probability robot is above landing zone?

$$p\left(x = \text{home} \mid z = \text{bright}\right)$$



Example adapted from Prob. Robotics book and Dr. J. Sturm

Example

Bayes rule

$$\boxed{p(x|z) = \frac{p(z|x)\,p(x)}{p(z)}}$$

- Binary sensor: $z \in \{\text{bright}, \neg\text{bright}\}$
- Binary state: $x \in \{\text{home}, \neg\text{home}\}$
- Prior on state: $p(x = \text{home}) = \underline{0.5}$
- Sensor model:

$$p(z = \text{bright} \mid x = \text{home}) = 0.6 \qquad p(z = \neg\text{bright} \mid x = \text{home}) = 0.4$$

$$p(z = \text{bright} \mid x = \neg\text{home}) = 0.3 \qquad p(z = \neg\text{bright} \mid x = \neg\text{home}) = 0.7$$

- Bayes rule:

$$p(x = \text{home} \mid z = \text{bright}) = \frac{p(z = \text{bright} \mid x = \text{home})\,p(x = \text{home})}{p(z = \text{bright})}$$

$$= \frac{p(z = \text{bright} \mid x = \text{home})\,p(x = \text{home})}{p(z = \text{bright} \mid x = \text{home})\,p(x = \text{home}) + p(z = \text{bright} \mid x = \neg\text{home})\,p(x = \neg\text{home})}$$

$$= \frac{0.6 \cdot 0.5}{0.6 \cdot 0.5 + 0.3 \cdot 0.5} = \frac{0.3}{0.3 + 0.15} = \underline{0.67}$$

Example adapted from Prob. Robotics book and Dr. J. Sturm

Combining Evidence

- Suppose the robot obtains another observation z_2 (from the same or a different sensor)
- How to integrate this new information?

- More generally:
  - How can we estimate $p\left(x|z_1, z_2, \ldots\right)$?
  - What if state and observations are continuous and not binary?
  - What about actions/controls?

↓

Bayesian Inference

Bayesian Updates

- Bayes rule, given some known quantity a: $\boxed{p\left(x|z,a\right)=\frac{p\left(z|x,a\right)p\left(x|a\right)}{p\left(z|a\right)}}$
- Therefore: $p\left(x|z_1,\ldots,z_n\right)=\frac{p\left(z_n|x,z_1,\ldots,z_{n-1}\right)p\left(x|z_1,\ldots,z_{n-1}\right)}{p\left(z_n|z_1,\ldots,z_{n-1}\right)}$
- Markov assumption: $p\left(z_n|x\right)=p\left(z_n|x,z_1,\ldots,z_{n-1}\right)$
  - x is sufficient for predicting observation

Bayesian Updates

- Bayes rule, given some known quantity a: $\boxed{p\left(x|z,a\right)=\frac{p\left(z|x,a\right)p\left(x|a\right)}{p\left(z|a\right)}}$

- Therefore: $p\left(x|z_1,\ldots,z_n\right)=\frac{p\left(z_n|x,z_1,\ldots,z_{n-1}\right)p\left(x|z_1,\ldots,z_{n-1}\right)}{p\left(z_n|z_1,\ldots,z_{n-1}\right)}$

- Markov assumption: $p\left(z_n|x\right)=p\left(z_n|x,z_1,\ldots,z_{n-1}\right)$

$$\Rightarrow \quad p\left(x|z_1,\ldots,z_n\right)=\frac{p\left(z_n|x\right)p\left(x|z_1,\ldots,z_{n-1}\right)}{p\left(z_n|z_1,\ldots,z_{n-1}\right)}$$

$$=\eta p\left(z_n|x\right)p\left(x|z_1,\ldots,z_{n-1}\right)$$

$$=\eta_{1:n}\prod_{i=1}^{n}p\left(z_i|x\right)p\left(x\right)$$

[Derivation @ whiteboard]

Example: Second Measurement

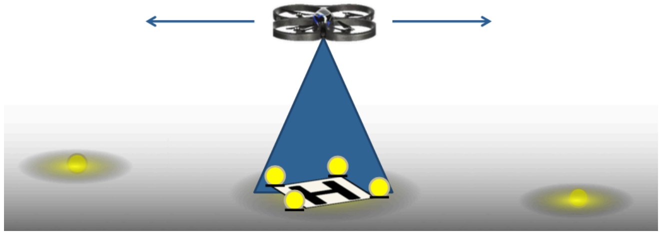
- Binary sensor: $z \in \{\text{bright}, \neg\text{bright}\}$
- Binary state: $x \in \{\text{home}, \neg\text{home}\}$
- Prior on state: $p\left(x = \text{home}\right) = 0.5$

- Sensor model: $p\left(z = \text{bright} \mid x = \text{home}\right) = 0.6$ $\quad p\left(z = \neg\text{bright} \mid x = \text{home}\right) = 0.4$

  $p\left(z = \text{bright} \mid x = \neg\text{home}\right) = 0.3$ $\quad p\left(z = \neg\text{bright} \mid x = \neg\text{home}\right) = 0.7$

- Assume now the robot makes a new observation: $z_2 = \neg\text{bright}$
- What is the probability robot is above landing zone (i.e. home)?

$$p\left(x = \text{home} \mid z_1 = \text{bright}, z_2 = \neg\text{bright}\right)$$



Example adapted from Prob. Robotics book and Dr. J. Sturm

Example: Second Measurement

$$p(x|z,a) = \frac{p(z|x,a)\,p(x|a)}{p(z|a)}$$

- Binary sensor: $z \in \{\text{bright}, \neg\text{bright}\}$
- Binary state: $x \in \{\text{home}, \neg\text{home}\}$
- Prior on state: $p(x = \text{home}) = 0.5$

- Sensor model: $p(z = \text{bright} \mid x = \text{home}) = 0.6$ $\quad p(z = \neg\text{bright} \mid x = \text{home}) = 0.4$

  $p(z = \text{bright} \mid x = \neg\text{home}) = 0.3$ $\quad p(z = \neg\text{bright} \mid x = \neg\text{home}) = 0.7$

- Previously (for the 1st measurement): $p(x = \text{home} \mid z_1 = \text{bright}) = 0.67$

- Therefore:

$$p(x = \text{home} \mid z_1 = \text{bright}, z_2 = \neg\text{bright}) =$$

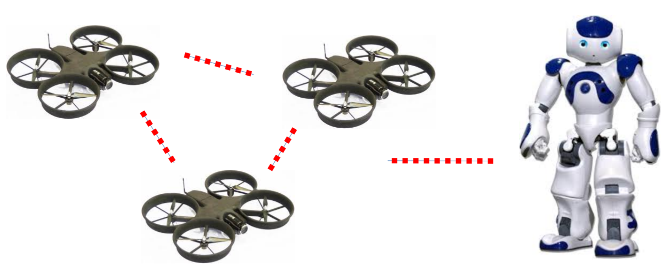
$$= \frac{p(z_2 = \neg\text{bright} \mid x = \text{home})\, p(x = \text{home} \mid z_1 = \text{bright})}{p(z_2 = \neg\text{bright}|x = \text{home})\, p(x = \text{home}|z_1 = \text{bright}) + p(z_2 = \neg\text{bright}|x = \neg\text{home})\, p(x = \neg\text{home}|z_1 = \text{brigh}}$$

$$= \frac{0.4 \cdot 0.67}{0.4 \cdot 0.67 + 0.7 \cdot 0.33} = 0.537 \downarrow$$

Example adapted from Prob. Robotics book and Dr. J. Sturm

Actions/Controls

- Examples for actions, controls?
  - Actions carried out by the robot
  - Actions carried out by other agents (robots, humans etc.)
  - Dynamically changing environment/world over time
- Actions are (almost) never carried out with absolute certainty
- In contrast to measurements, actions generally <u>increase</u> uncertainty



Multiple robots

CMU

Actions/Controls

- How can we incorporate actions into Bayesian inference?

- To incorporate the outcome of an action *u* into the current state estimate ("belief"), we use the conditional pdf

$$p\left(x' | x, u\right)$$

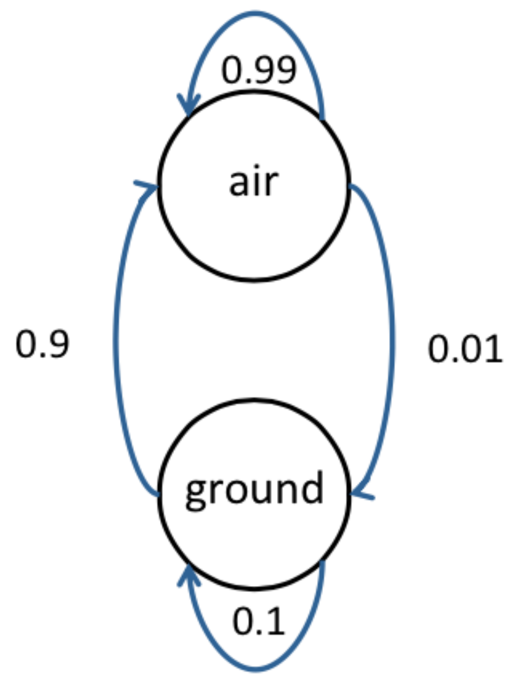
  – Specifies the probability that executing action *u* in state *x* will lead to state *x'*
  – Also known as motion model and state transition model (formal definition - soon)

Incorporating the Outcome of Actions

- Recall – chain rule: $p(x, y) = p(x|y)\, p(y)$
- Therefore: $p(x', x|u) = p(x'|x, u)\, p(x)$
- If only the new state is of interest, need to integrate out (marginalize) previous state:
  - Discrete state space: $$p(x'|u) = \sum_x p(x'|x, u)\, p(x)$$
  - Continuous state space: $$p(x'|u) = \int p(x'|x, u)\, p(x)\, dx$$
- What happens without marginalization?

Example: Take-Off

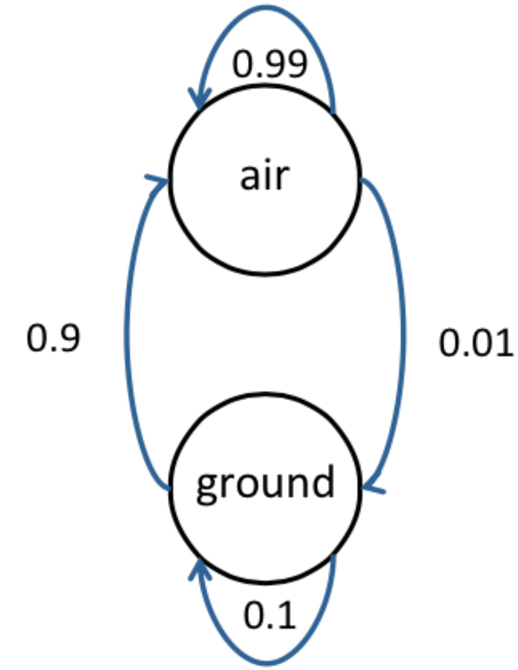
- Action: $u \in \{\text{takeoff}\}$
- State: $x \in \{\text{ground}, \text{air}\}$



Example by Dr. J. Sturm

Example

- Prior belief on robot state: $p\left(x=\text{ground}\right)=1$
- Robot/Operator issues a “take-off” action
- Robot’s belief after one time step:



$$p\left(x'=\text{ground}|u\right)=\sum_{x}p\left(x'=\text{ground}|u,x\right)p\left(x\right)$$

$$=p\left(x'=\text{ground}|u,x=\text{ground}\right)p\left(x=\text{ground}\right)+p\left(x'=\text{ground}|u,x=\text{air}\right)p\left(x=\text{air}\right)$$

$$=0.1\cdot 1+0.01\cdot 0=0.1$$

$$\boxed{p\left(x'|u\right)=\sum_{x}p\left(x'|x,u\right)p\left(x\right)}$$

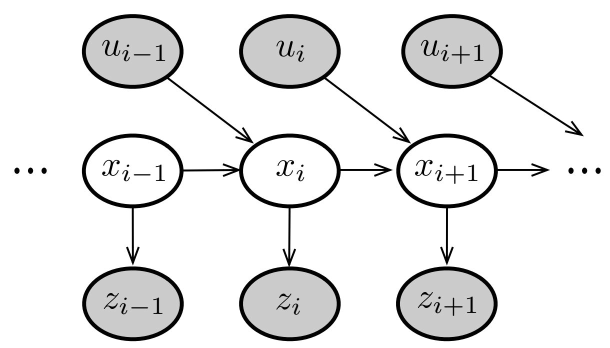
Example by Dr. J. Sturm

Formal Exposition

Bayesian Inference

- **Formally**, Bayesian inference is an approach in which Bayes rule is used to update the probability over variables of interest with new data (sensor readings, actions)

- Generative Model - Markov Chain:
  – A Markov chain is a stochastic process where, given the present state, past and future states are independent



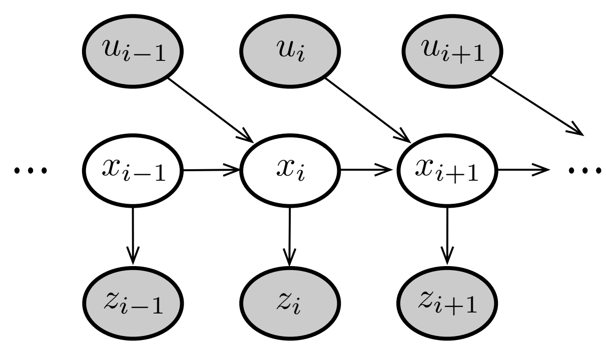
Bayesian Inference

- Motion model (or state transition model)

$$x_{k+1} = f\left(x_k, u_k\right) + w_k \qquad\qquad p\left(x_{k+1} | x_k, u_k\right)$$

- Observation model

$$z_k = h\left(x_k\right) + v_k \qquad\qquad p\left(z_k | x_k\right)$$



Bayesian Inference

- Motion model (or state transition model)

$$x_{k+1} = f(x_k, u_k) + w_k \qquad\qquad p(x_{k+1} | x_k, u_k)$$

- Observation model

$$z_k = h(x_k) + v_k \qquad\qquad p(z_k | x_k)$$

- A posteriori pdf given **all** measurements and controls:

$$p(x_k | u_{0:k-1}, z_{1:k})$$

- Objective - Maximum a posteriori (MAP) estimation:

$$x_k^{\star} = \arg\max_{x_k} p(x_k | u_{0:k-1}, z_{1:k})$$

Bayesian Inference

- Markov assumption

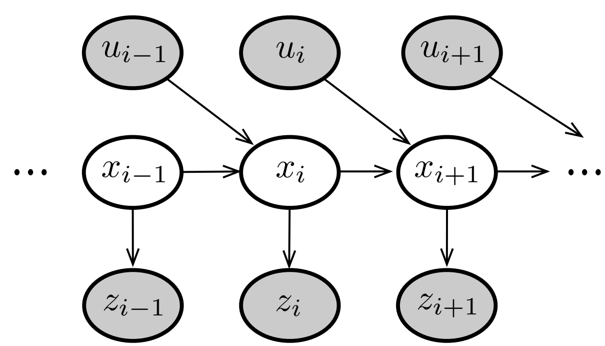
$$p\left(x_k | x_{0:k-1}, z_{1:k-1}, u_{0:k-1}\right) = p\left(x_k | x_{k-1}, u_{k-1}\right)$$

$$p\left(z_k | x_{0:k}, u_{0:k-1}, z_{1:k-1}\right) = p\left(z_k | x_k\right)$$

  – Corresponds to assuming independent noise:

$$\forall i \neq j : \quad \mathbb{E}\left[w_i w_j^T\right] = 0 \ , \ \mathbb{E}\left[v_i v_j^T\right] = 0$$

$$\forall i, j : \qquad \mathbb{E}\left[v_i w_j^T\right] = 0$$



Recursive Bayesian Update

Recall:

[Derivation @ whiteboard]

- Motion model $x_{k+1} = f(x_k, u_k) + w_k$ $\quad p(x_{k+1} | x_k, u_k)$
- Observation model $z_k = h(x_k) + v_k$ $\quad p(z_k | x_k)$

- Given a posteriori pdf from previous time $p(x_{k-1} | u_{0:k-2}, z_{1:k-1})$
- Objective: calculate $p(x_k | u_{0:k-1}, z_{1:k})$

Calculation can be described in two steps:

- Prediction

$$p(x_k | u_{0:k-1}, z_{1:k-1}) = \int \underbrace{p(x_{k-1} | u_{0:k-2}, z_{1:k-1})}_{\textbf{previous belief}} p(x_k | x_{k-1}, u_{k-1})\, dx_{k-1}$$

- Update $p(x_k | u_{0:k-1}, z_{1:k}) = \eta p(z_k | x_k) p(x_k | u_{0:k-1}, z_{1:k-1})$

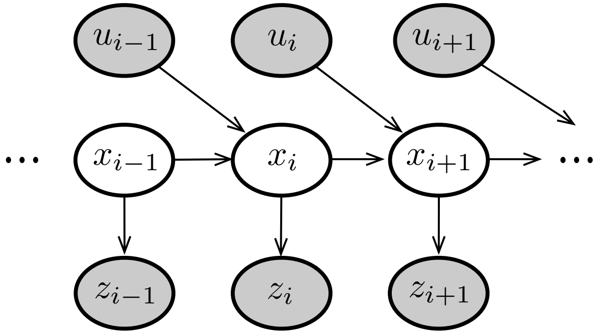
Recursive Estimation vs Smoothing

- What if we do not marginalize past states?

$$p\left(x_{0:k}|u_{0:k-1}, z_{1:k}\right) = \eta p\left(x_0\right) \prod_{i=1}^{k} p\left(x_i|x_{i-1}, u_{i-1}\right) p\left(z_i|x_i\right)$$

  – Smoother: past states are updated with future observations
    - Not causal
    - Important in many applications (e.g. mapping)
    - Commonly used in SLAM and SfM (more details soon)

  – Fixed-lag smoother



Bayesian Inference - Summary

- Markov assumption allows efficient recursive Bayesian updates of the belief distribution

- Bayesian inference (Bayes filter) is the basis of many filters
  - (Extended) Kalman filter - **next**
  - Particle filter
  - Hidden Markov models

- Used also in planning and decision making under uncertainty
  - Partially Observable Markov Decision Process (POMDP)
  - Belief space planning

(Extended) Kalman Filter

[Derivation @ whiteboard]

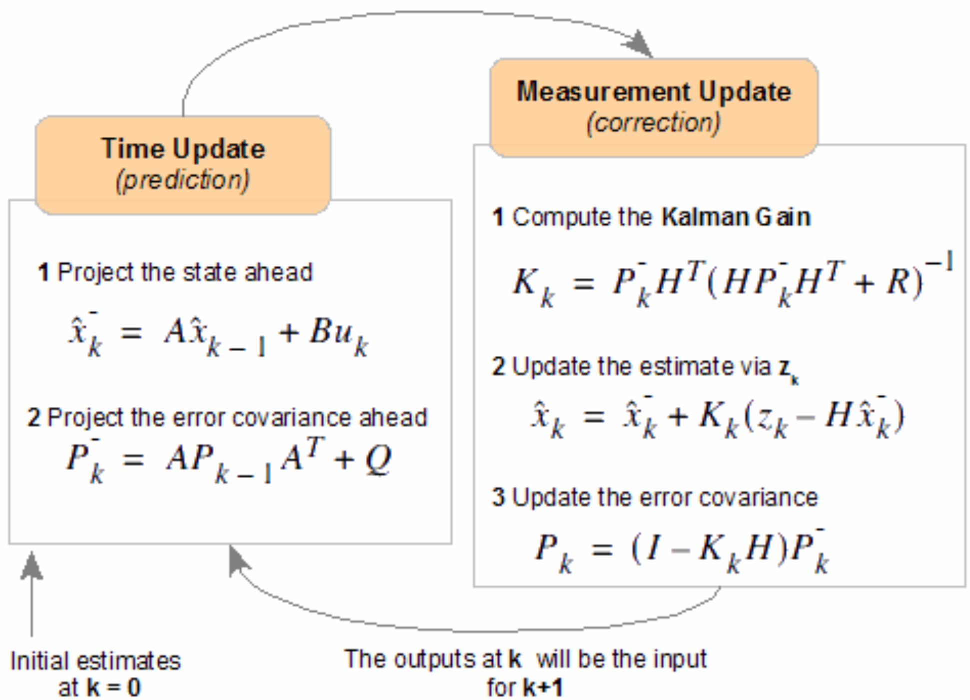
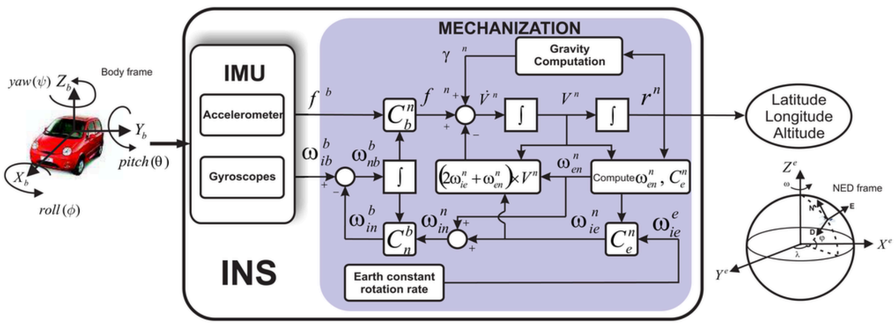
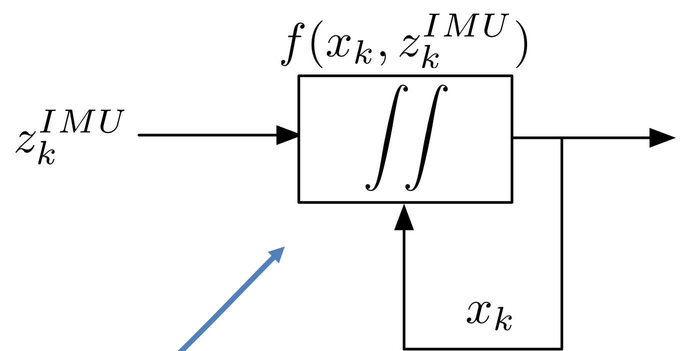


Image from http://bilgin.esme.org/BitsBytes/KalmanFilterforDummies.aspx

Example: INS/GPS System

- Navigation state: $x \doteq \begin{bmatrix} p^T & v^T & \Psi^T & b_a^T & b_g^T \end{bmatrix}^T \in \mathbb{R}^{15}$
- Motion model: INS mechanization equations





- Jacobian: $F = \nabla_x f = \dfrac{\partial f(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}} = \ldots$

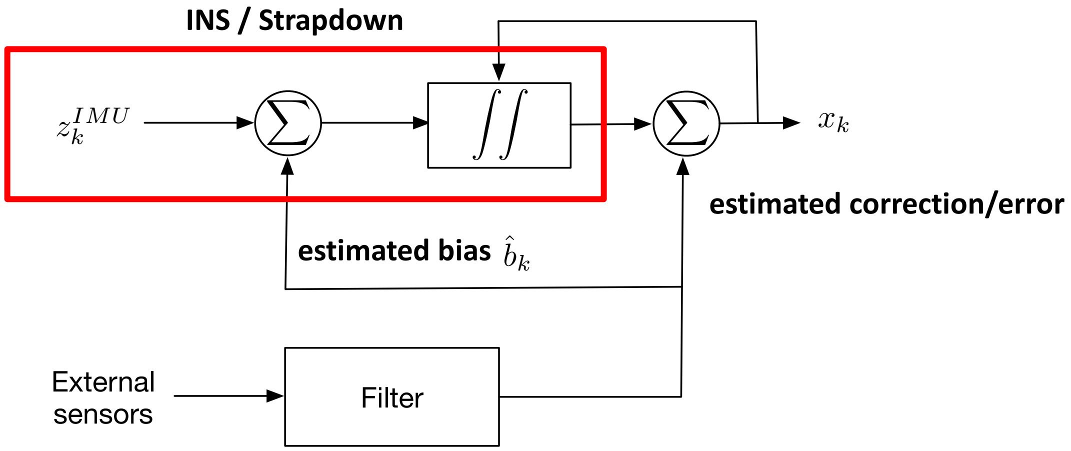
Example: INS/GPS System

- Observation function (GPS): $z_k = h\left(x_k\right) + v_k$

$$h\left(\mathbf{x}\right) = \ldots$$

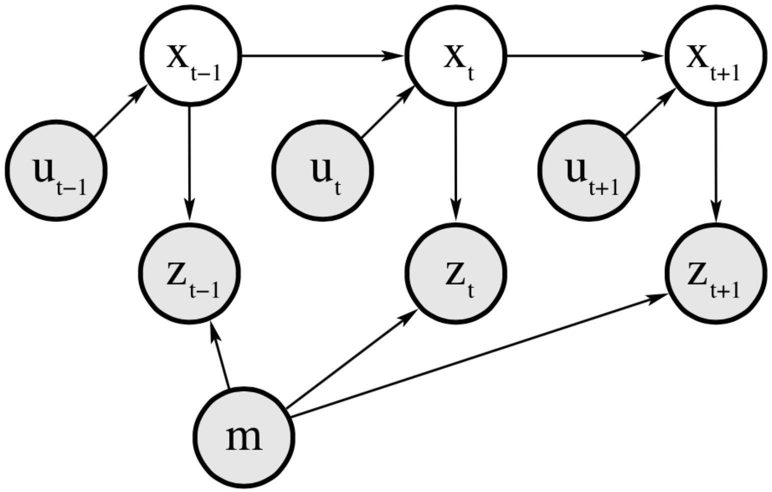
$$H = \nabla_x h = \frac{\partial h\left(\mathbf{x}\right)}{\partial \mathbf{x}} = \ldots$$

- Different rates (IMU – high-frequency)
  - Perform propagate step until GPS measurement is received
  - Navigation-aiding scheme



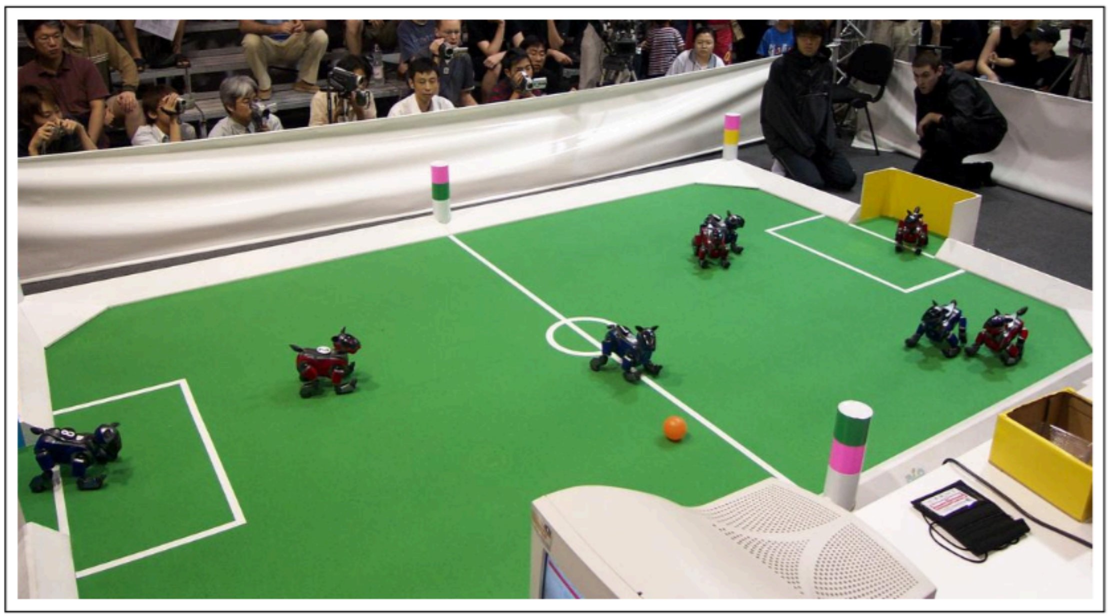
Example: Mobile Robot Localization

- Shaded nodes are known parameters:
  - A given map m
  - Controls u
  - Measurements (e.g. laser range sensor, camera)
- Objective: Infer robot state x
- Basic equations are the same



Example: Mobile Robot Localization

- AIBO robots on the RoboCup soccer field
- Six **known** landmarks are placed at the corners and the midlines of the field.



Images from Probabilistic Robotics book, chapter 7.4

Example: Mobile Robot Localization

- Different types of map representation
  - Hand-drawn
  - Topological
  - Occupancy grid
  - Image mosaic

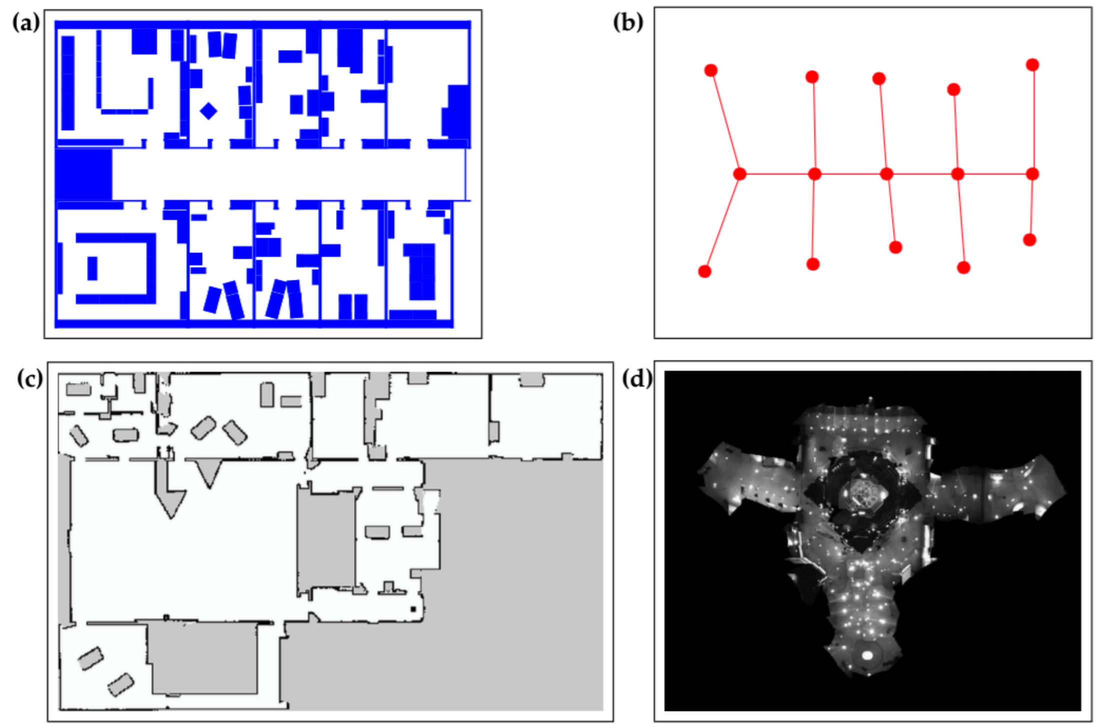


Figure 7.2 Example maps used for robot localization: (a) a manually constructed 2-D metric layout, (b) a graph-like topological map, (c) an occupancy grid map, and (d) an image mosaic of a ceiling. (d) courtesy of Frank Dellaert, Georgia Institute of Technology.

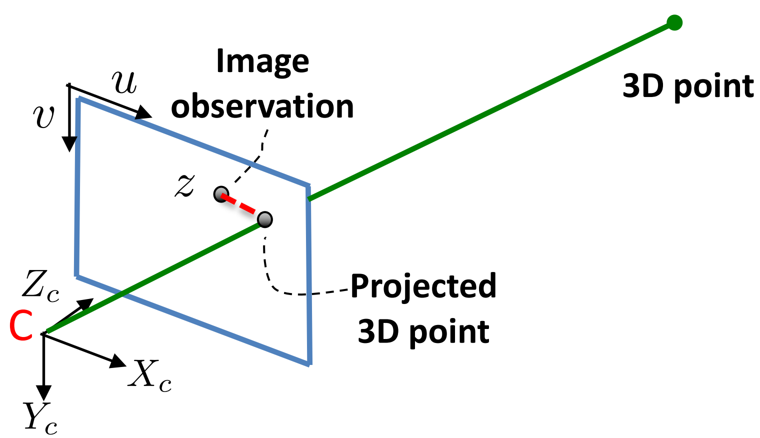
Images from Probabilistic Robotics book, chapter 7.1

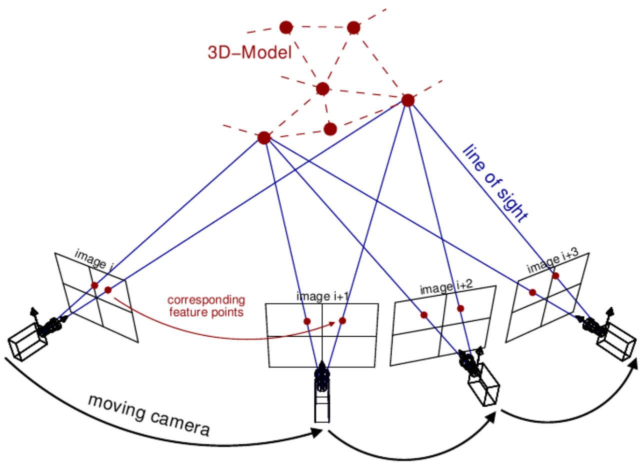
Recap – Probabilistic Inference

- Approach to update probability distribution function with new information (e.g. sensor measurements) or events (e.g. controls, actions)
- Recursive vs smoothing formulation
- Extended Kalman filter

Next

- Projective Cameras
- What if map (or 3D points) is **<u>not</u>** given, or inaccurate?





References

- Probabilistic Robotics, Chapters 2 and 3