Probabilistic Robotics Course

Data Association

G. Grisetti
grisetti@diag.uniroma1.it

Department of Computer, Control and Management Engineering
Sapienza University of Rome

Outline

- Problem Definition
- Finding the best association
- Computing the prior of the measurements
- Gaussian case
- Evaluating an association
- (En)forcing independence
- Multiple assignments
- Greedy Heuristics

Scenario

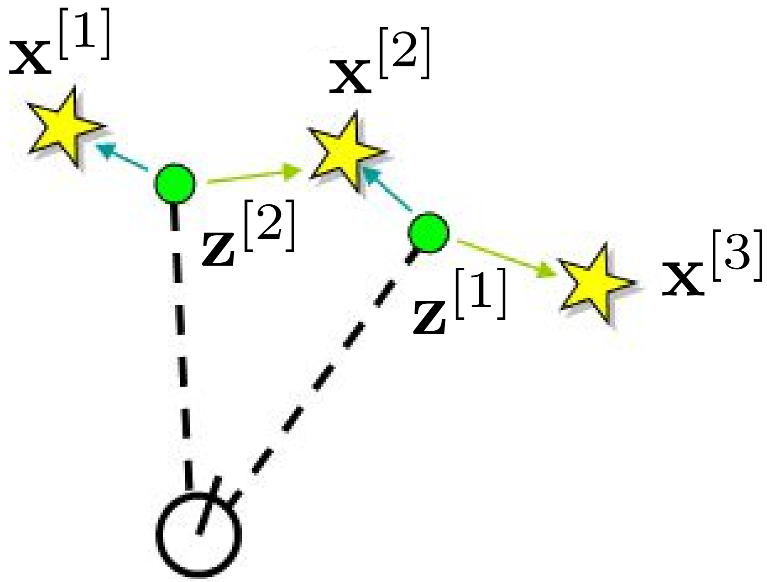
In the problems we approached the landmarks were uniquely identifiable

More formally we assumed to know which observation was generated by which landmark

Bad things happen if this is not the case

Our inference engine will have to deal with distributions having multiple maxima	→	Deal with multi-modal distributions No EKF/UKF
Choose the best possible association	→	Approximations that might lessen robustness

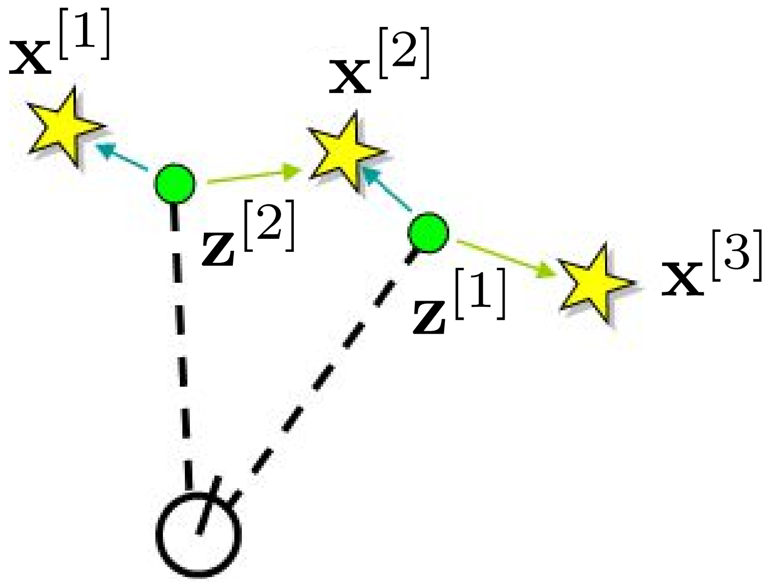
Data Association



Determine which state variable (if any) is responsible of a measurement, based on:

- the current estimate of the state
- the history of measurements

Data Association



How many possible combinations?

If we have *n* landmarks and *m* measurements, with *n*>*m* the number of possible associations is

$$\frac{n!}{(n-m)!} = \left(\begin{array}{c} n \\ m \end{array} \right)$$

Best Association

The best association j^* maximizes the joint likelihood of the measurements:

$$\mathrm{j}^* = \underset{j}{\operatorname{argmax}}\, p(\hat{\mathbf{z}}^{[\mathrm{j}(1)]} = \mathbf{z}^{[1]}, \ldots, \hat{\mathbf{z}}^{[\mathrm{j}(M)]} = \mathbf{z}^{[M]})$$

probability of assignment **j**

$$\mathrm{j}(i) \in [0, 1, 2, \ldots N]$$

j(3)=5 means measurement 3 originates from landmark 5
j(n)=0 means no correspondence

$$\mathbf{x} = \mathbf{x}^r, \mathbf{x}^{[1]}, \ldots, \mathbf{x}^{[N]}$$

state (robot, landmarks)

$$\mathbf{z} = \mathbf{z}^{[1]}, \ldots, \mathbf{z}^{[M]}$$

measurements (1, .., M)

$$\hat{\mathbf{z}} = \hat{\mathbf{z}}^{[1]}, \ldots, \hat{\mathbf{z}}^{[N]}$$

predictions (1, .., N)

Evaluating an Association

We can write down the steps to compute the likelihood of an association ***j***

chain rule + marginalization

$$p(\hat{\mathbf{z}}^{[\mathrm{j}(1)]} = \mathbf{z}^{[1]}, \ldots, \hat{\mathbf{z}}^{[\mathrm{j}(M)]} = \mathbf{z}^{[M]})$$

$$= \int_{\mathbf{x}} p(\hat{\mathbf{z}}^{[\mathrm{j}(1)]} = \mathbf{z}^{[1]}, \ldots, \hat{\mathbf{z}}^{[\mathrm{j}(M)]} = \mathbf{z}^{[M]} \mid \mathbf{x}) p(\mathbf{x})$$

$$= \int_{\mathbf{x}\in\Omega} \prod_{m=1}^{M} p(\mathbf{z}^{[\mathrm{j}(m)]} = \mathbf{z}^{[m]} \mid \mathbf{x}^{[\mathrm{j}(m)]}, \mathbf{x}^{r}) p(\mathbf{x}^{[\mathrm{j}(m)]}, \mathbf{x}^{r})$$

measurement independence

We have to repeat this for each association

Can we do something better?

Joint Prediction

We notice that the distribution of the prediction is not dependent on the specific assignment

$$p(\hat{\mathbf{z}}) = \int_{\mathbf{x}} p(\hat{\mathbf{z}} \mid \mathbf{x}) p(\mathbf{x})$$

chain rule + marginalization

$$= \int_{\mathbf{x}\in\Omega} \prod_{n=1}^{N} p(\hat{\mathbf{z}}^{[n]} \mid \mathbf{x}^{[n]}, \mathbf{x}^{r}) p(\mathbf{x})$$

measurement independence

This has N blocks (as many as the landmarks)

We can compute this distribution **once**, and then evaluate it for each assignment

Independence Loss

Marginalizing out the states renders the landmark observations not independent, thus we cannot "estimate" the likelihood of a single landmark assignment as the product of the single likelihoods

$$p(\hat{\mathbf{z}} = \mathbf{z}^{[\mathrm{j}]}) \neq \prod_{n=1}^{N} p(\hat{\mathbf{z}}^{[n]} = \mathbf{z}^{[\mathrm{j}^{-1}(n)]})$$

this means NOT equal

Inverse Assignment

Having defined a mapping from landmarks to measurements, we can also define the inverse mapping from measurements to landmarks

$$\mathrm{j}(m) = n$$

$$\mathrm{j}^{-1}(n) = m$$

Some landmarks might not appear in the observations. In this case j(n) will return an invalid value.

The mapping is not invertible!

Using the Joint Prediction

Having the prediction of all landmarks, to determine how likely is one specific assignment we can

$$\mathbf{z}^{[\mathrm{j}]} = \begin{pmatrix} \mathbf{z}^{[\mathrm{j}^{-1}(1)]} \\ \mathbf{z}^{[\mathrm{j}^{-1}(2)]} \\ \mathbf{z}^{[\mathrm{j}^{-1}(3)]} \\ \vdots \\ \mathbf{z}^{[\mathrm{j}^{-1}(N)]} \end{pmatrix}$$

Reshuffle the measurement vector by reordering the measures according to the inverse assignment $\mathbf{j}^{-1}$

Using the Joint Prediction

Having the prediction of all landmarks, to determine how likely is one specific assignment we can

$$\mathbf{z}^{[\mathrm{j}]} = \begin{pmatrix} \mathbf{z}^{[\mathrm{j}^{-1}(1)]} \\ \\ \mathbf{z}^{[\mathrm{j}^{-1}(3)]} \\ \vdots \\ \mathbf{z}^{[\mathrm{j}^{-1}(N)]} \end{pmatrix}$$

Reshuffle the measurement vector by reordering the measures according to the inverse assignment $\mathbf{j}^{-1}$

The mapping is not invertible, so we will have "holes" for landmarks not seen or unmatched measurements

Gaussian Case

We have our usual Gaussian assumptions, and we apply marginalization and chain rule in one go

$$p(\mathbf{z} \mid \mathbf{x}) = \mathcal{N}(\mathbf{z}; \mathbf{C}\mathbf{x} + \mathbf{d}, \mathbf{\Sigma}_{z|x})$$

conditional over all landmarks

$$p(\mathbf{x}) = \mathcal{N}(\mathbf{x}; \mu_x, \mathbf{\Sigma}_x)$$

prior on pose and landmarks

$$\mu = \begin{pmatrix} \mu^r \\ \mu^{[1]} \\ \mu^{[2]} \\ \vdots \\ \mu^{[N]} \end{pmatrix} \qquad \mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}^{[rr]} & \mathbf{\Sigma}^{[r1]} & \mathbf{\Sigma}^{[r2]} & \dots & \mathbf{\Sigma}^{[rN]} \\ \mathbf{\Sigma}^{[1r]} & \mathbf{\Sigma}^{[11]} & \mathbf{\Sigma}^{[12]} & \dots & \mathbf{\Sigma}^{[1N]} \\ \mathbf{\Sigma}^{[2r]} & \mathbf{\Sigma}^{[21]} & \mathbf{\Sigma}^{[22]} & \dots & \mathbf{\Sigma}^{[2N]} \\ \vdots & & & & \\ \mathbf{\Sigma}^{[Nr]} & \mathbf{\Sigma}^{[N1]} & \mathbf{\Sigma}^{[N2]} & \dots & \mathbf{\Sigma}^{[NN]} \end{pmatrix}$$

$$\mu_z = \mathbf{C}\mu_x + \mathbf{d} \qquad \mathbf{\Sigma}_z = \mathbf{C}\mathbf{\Sigma}_x\mathbf{C}^T + \mathbf{\Sigma}_{z|x} = \begin{pmatrix} \mathbf{\Sigma}_z^{[11]} & \mathbf{\Sigma}_z^{[12]} & \dots & \mathbf{\Sigma}_z^{[1N]} \\ \mathbf{\Sigma}_z^{[21]} & \mathbf{\Sigma}_z^{[22]} & \dots & \mathbf{\Sigma}_z^{[2N]} \\ \vdots & & & \\ \mathbf{\Sigma}_z^{[N1]} & \mathbf{\Sigma}_z^{[N2]} & \dots & \mathbf{\Sigma}_z^{[NN]} \end{pmatrix}$$

joint covariance of the landmarks

Gaussian Case

Evaluating an association

hole

$$\mathbf{z}^{[\mathrm{j}]} = \begin{pmatrix} \mathbf{z}^{[\mathrm{j}^{-1}(1)]} \\ \boxed{\qquad\qquad} \\ \mathbf{z}^{[\mathrm{j}^{-1}(3)]} \\ \vdots \\ \mathbf{z}^{[\mathrm{j}^{-1}(n)]} \\ \vdots \\ \mathbf{z}^{[\mathrm{j}^{-1}(N)]} \end{pmatrix}$$

Reshuffle the measurement vector by reordering the measures according to the inverse assignment $\mathbf{j}^{-1}$

Gaussian Case

Evaluating an association

$$\mathbf{z}^{[\mathrm{j}]} = \begin{pmatrix} \mathbf{z}^{[\mathrm{j}^{-1}(1)]} \\ \\ \mathbf{z}^{[\mathrm{j}^{-1}(3)]} \\ \vdots \\ \mathbf{z}^{[\mathrm{j}^{-1}(n)]} \\ \vdots \\ \mathbf{z}^{[\mathrm{j}^{-1}(N)]} \end{pmatrix}$$

Suppress all "holes" in
- measurements
- prediction
- covariance matrix

  This corresponds to marginalizing

$$\mu_z = \begin{pmatrix} \mu_z^{[1]} \\ \\ \vdots \\ \mu_z^{[N]} \end{pmatrix} \qquad \mathbf{\Sigma}_z = \begin{pmatrix} \mathbf{\Sigma}_z^{[11]} & & \cdots & \mathbf{\Sigma}_z^{[1N]} \\ & & & \\ \vdots & & & \\ \mathbf{\Sigma}_z^{[N1]} & & \cdots & \mathbf{\Sigma}_z^{[NN]} \end{pmatrix}$$

Gaussian Case

Evaluating an association

$$\tilde{\mathbf{z}}^{[\mathrm{j}]} = \begin{pmatrix} \mathbf{z}^{[\mathrm{j}^{-1}(1)]} \\ \mathbf{z}^{[\mathrm{j}^{-1}(3)]} \\ \vdots \\ \mathbf{z}^{[\mathrm{j}^{-1}(n)]} \\ \vdots \\ \mathbf{z}^{[\mathrm{j}^{-1}(N)]} \end{pmatrix}$$

Suppress all “holes” in

- measurements
- prediction
- covariance matrix

This corresponds to marginalizing

$$\tilde{\mu}_z = \begin{pmatrix} \mu_z^{[1]} \\ \vdots \\ \mu_z^{[N]} \end{pmatrix} \qquad \tilde{\mathbf{\Sigma}}_z = \begin{pmatrix} \mathbf{\Sigma}_z^{[11]} & \cdots & \mathbf{\Sigma}_z^{[1N]} \\ \vdots & & \\ \mathbf{\Sigma}_z^{[N1]} & \cdots & \mathbf{\Sigma}_z^{[NN]} \end{pmatrix}$$

Gaussian Case

Evaluating an association

$$\tilde{\mathbf{z}}^{[\mathrm{j}]} = \begin{pmatrix} \mathbf{z}^{[\mathrm{j}^{-1}(1)]} \\ \mathbf{z}^{[\mathrm{j}^{-1}(3)]} \\ \vdots \\ \mathbf{z}^{[\mathrm{j}^{-1}(n)]} \\ \vdots \\ \mathbf{z}^{[\mathrm{j}^{-1}(N)]} \end{pmatrix}$$

Evaluate the Gaussian

$$p(\hat{\mathbf{z}}^{[\mathrm{j}(1)]} = \mathbf{z}^{[1]}, \ldots, \hat{\mathbf{z}}^{[\mathrm{j}(M)]} = \mathbf{z}^{[M]})$$

$$\propto \frac{1}{\ldots} \exp\left(-(\tilde{\mathbf{z}}^{[j]} - \tilde{\mu}_z)^T \tilde{\Sigma}_z^{-1} (\tilde{\mathbf{z}}^{[j]} - \tilde{\mu}_z) \right)$$

$$\tilde{\mu}_z = \begin{pmatrix} \mu_z^{[1]} \\ \vdots \\ \mu_z^{[N]} \end{pmatrix} \qquad \tilde{\mathbf{\Sigma}}_z = \begin{pmatrix} \mathbf{\Sigma}_{11} & \cdots & \mathbf{\Sigma}_{1N} \\ \vdots & & \\ \mathbf{\Sigma}_{N1} & \cdots & \mathbf{\Sigma}_{NN} \end{pmatrix}$$

Beware the tilde

Holes

The above procedure has an issue

It favors associations with zero or very few landmarks

To account for this we multiply the resulting value with a factor **p_{miss}**<1 exponential in the number of holes

$$p(\hat{\mathbf{z}}^{[\mathrm{j}(1)]} = \mathbf{z}^{[1]}, \ldots, \hat{\mathbf{z}}^{[\mathrm{j}(M)]} = \mathbf{z}^{[M]})$$

$$\propto \frac{1}{\ldots} \exp\left(-(\tilde{\mathbf{z}}^{[j]} - \tilde{\mu}_z)^T \tilde{\Sigma}_z^{-1} (\tilde{\mathbf{z}}^{[j]} - \tilde{\mu}_z)\right) \boxed{p_{\text{miss}}^K}$$

This number decreases with the number of unmatched measurements K

Independence Again

This renders the predictions independent

$$\mu_z = \begin{pmatrix} \mu_z^{[1]} \\ \mu_z^{[2]} \\ \vdots \\ \mu_z^{[N]} \end{pmatrix} \qquad \mathbf{\Sigma}_z = \begin{pmatrix} \mathbf{\Sigma}_z^{[11]} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_z^{[22]} & \cdots & \mathbf{0} \\ \vdots & & & \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Sigma}_z^{[NN]} \end{pmatrix}$$

$$p(\hat{\mathbf{z}} = \mathbf{z}^{[\mathrm{j}]}) = \prod_{m=1}^{M} p(\mathbf{z}^{[n]} = \hat{\mathbf{z}}^{[\mathrm{j}(n)]})$$

Maximizing this

$$\log p(\hat{\mathbf{z}} = \mathbf{z}^{[\mathrm{j}]}) = \sum_{m=1}^{M} \log p(\mathbf{z}^{[m]} = \hat{\mathbf{z}}^{[\mathrm{j}(m)]})$$

means minimizing this

$$\propto -\sum_{m=1}^{N} (\mathbf{z}^{[m]} - \mu_z^{[\mathrm{j}(m)]})^T \left(\mathbf{\Sigma}^{[j(m),j(m)]}\right)^{-1} (\mathbf{z}^{[m]} - \mu_z^{[\mathrm{j}(m)]})$$

High Complexity

The non independence of the measurements requires to evaluate the entire measurement vector as a "whole"

As a first approximation we might ignore the off diagonal blocks of the covariance matrix

$$\mu_z = \begin{pmatrix} \mu_z^{[1]} \\ \mu_z^{[2]} \\ \vdots \\ \mu_z^{[N]} \end{pmatrix} \qquad \mathbf{\Sigma}_z = \begin{pmatrix} \mathbf{\Sigma}_z^{[11]} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_z^{[22]} & \cdots & \mathbf{0} \\ \vdots & & & \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Sigma}_z^{[NN]} \end{pmatrix}$$

Hint: If you ignore these entries, no need to compute them

Independence Again

Each measurement can be associated with one of the N landmarks

The contribution to the log likelihood is additive

We can precompute the likelihood for each landmark/measurement pair

$$a_{mn} = (\mathbf{z}^{[m]} - \mu_z^{[n]})^T \mathbf{\Omega}^{[m,m]} (\mathbf{z}^{[m]} - \mu_z^{[n]})$$

computing the likelihood of an association

$$\sum_m a_{m,j(m)}$$

Independence and Assignment

Assemble these costs in a matrix:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots a_{1N} \\ a_{21} & a_{22} & a_{23} & \cdots a_{2N} \\ \vdots & & & \\ a_{M1} & a_{M2} & a_{M3} & \cdots a_{MN} \end{pmatrix}$$

Wiki for “assignment problem” and you will find this:

“The problem instance has a number of agents and a number of tasks. Any agent can be assigned to perform any task, incurring some cost that may vary depending on the agent-task assignment. It is required to perform all tasks by assigning exactly one agent to each task and exactly one task to each agent in such a way that the total cost of the assignment is minimized.”

Independence and Assignment

We want to find exactly one element for each column, such that

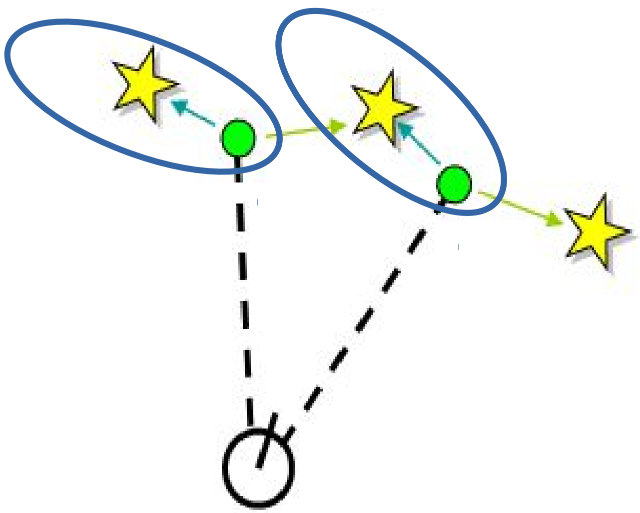
- the sum of all entries is minimized
- assignment: no column is selected more than once
- the matrix is made square by padding with a default value $\mathbf{l}_{\mathbf{miss}}$ modeling the cost of a miss

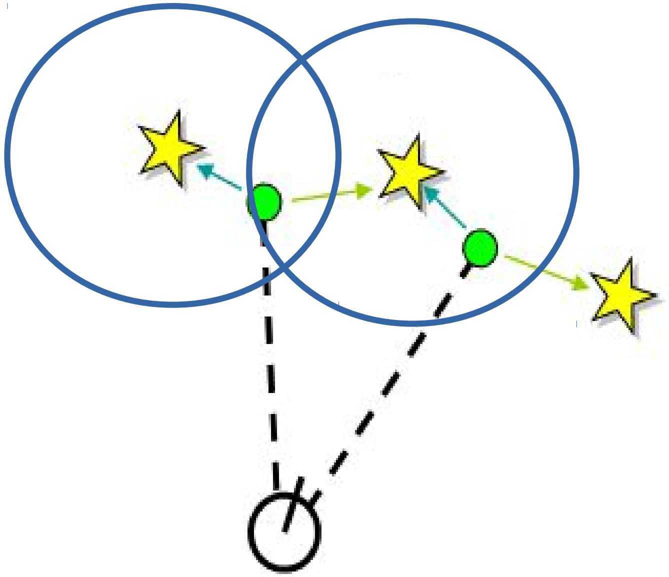
$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots a_{1N} \\ a_{21} & a_{22} & a_{23} & \cdots a_{2N} \\ \vdots & & & \\ a_{M1} & a_{M2} & a_{M3} & \cdots a_{MN} \\ l_{\text{miss}} & l_{\text{miss}} & l_{\text{miss}} & \cdots l_{\text{miss}} \\ l_{\text{miss}} & l_{\text{miss}} & l_{\text{miss}} & \cdots l_{\text{miss}} \end{pmatrix}$$

Nearest Neighbor

If the the measurement covariances have equal eigenvalues, the greedy association degenerates to the nearest neighbor strategy

Assign to each landmark the closest measurement





The nearest neighbor does not take into account the “shape” of the covariance

Nearest Neighbor

In case of nearest neighbor, the entries of the association matrix are the squared distances between prediction and measurement

$$\boxed{a_{mn}} = (\mathbf{z}^{[m]} - \mu_z^{[n]})^T (\mathbf{z}^{[m]} - \mu_z^{[n]})$$

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots a_{1N} \\ a_{21} & a_{22} & a_{23} & \cdots a_{2N} \\ \vdots & & & \\ a_{M1} & a_{M2} & a_{M3} & \cdots a_{MN} \\ l_{\text{miss}} & l_{\text{miss}} & l_{\text{miss}} & \cdots l_{\text{miss}} \\ l_{\text{miss}} & l_{\text{miss}} & l_{\text{miss}} & \cdots l_{\text{miss}} \end{pmatrix}$$

Even in the simple case of nearest neighbor, computing the association matrix can be expensive for large M,N. We will see efficient techniques on how to do this later in this course

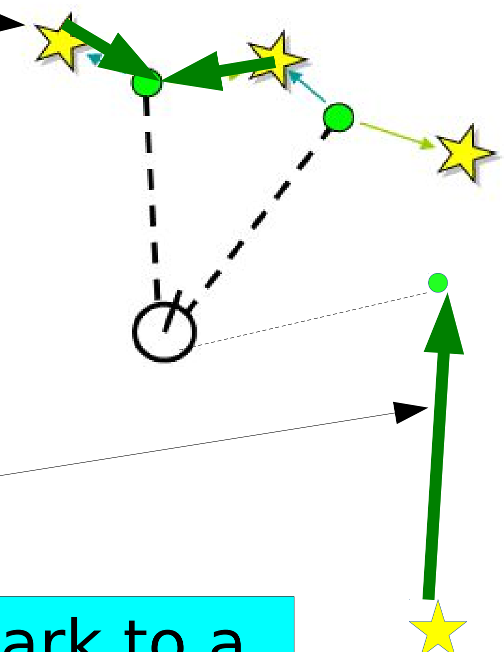
Greedy Association

A simpler solution consists in selecting for each column the row with the smaller value

Doing so, however might result in undesired configurations

Multiple measurements can be assigned to the same landmark

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots a_{1N} \\ a_{21} & a_{22} & a_{23} & \cdots a_{2N} \\ \vdots & & & \\ a_{M1} & a_{M2} & a_{M3} & \cdots a_{MN} \end{pmatrix}$$



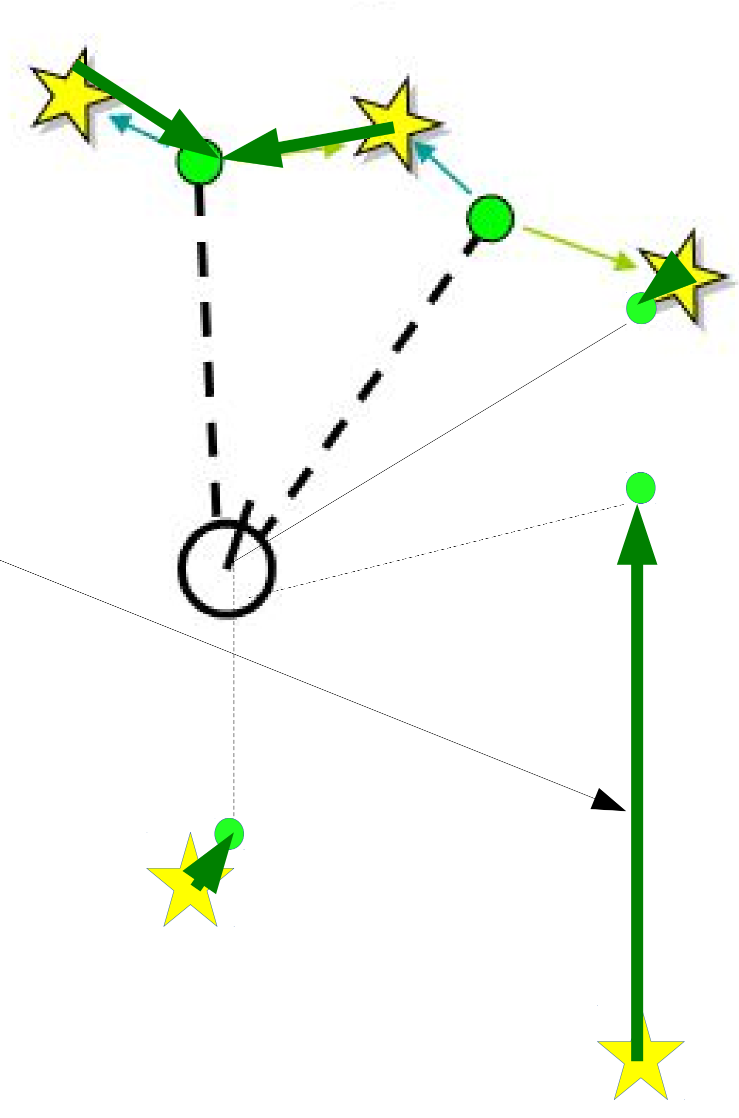
We can assign a landmark to a far measurement because nothing better is found

Pruning Heuristics

Bad associations are EVIL

To avoid the above cases we can use three heuristics

- Gating: ignore all associations whose cost is higher than a threshold
- Best friends: an association should be the best of both row and column
- Lonely best friends: the measurement should be clearly assigned to a single landmark and viceversa.

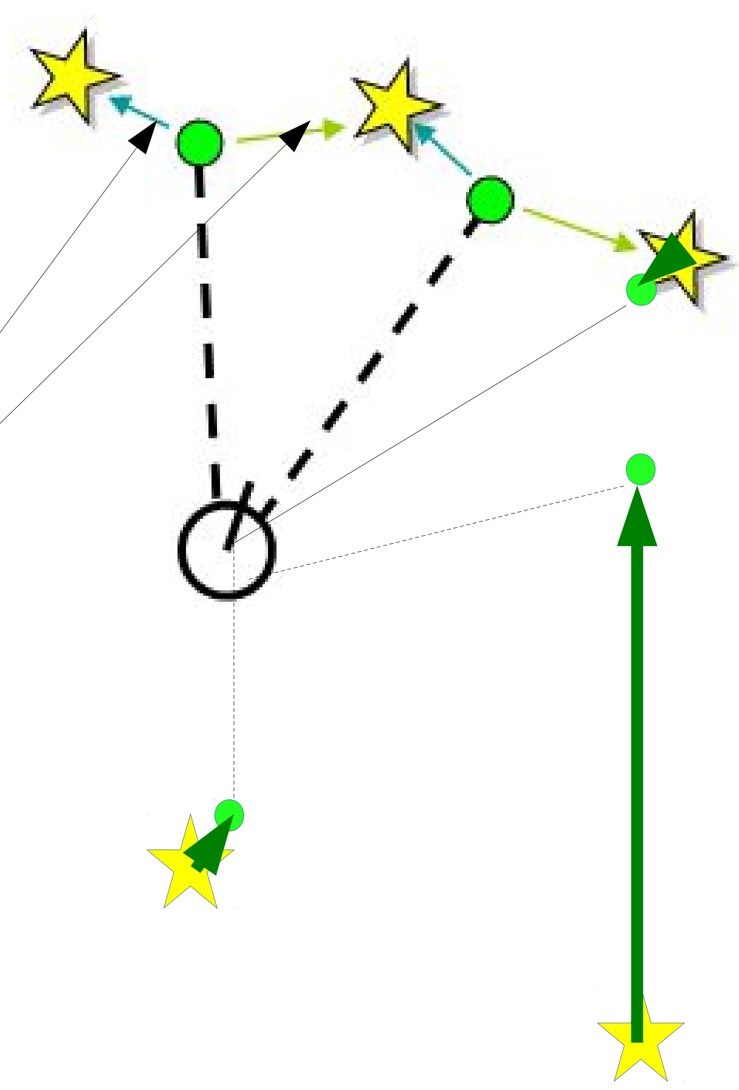


Pruning Heuristics

Bad associations are EVIL

To avoid the above cases we can use three heuristics

- Gating: ignore all associations whose cost is higher than a threshold
- Best friends: an association should be the best of both row and column
- Lonely best friends: the measurement should be clearly assigned to a single landmark and viceversa.

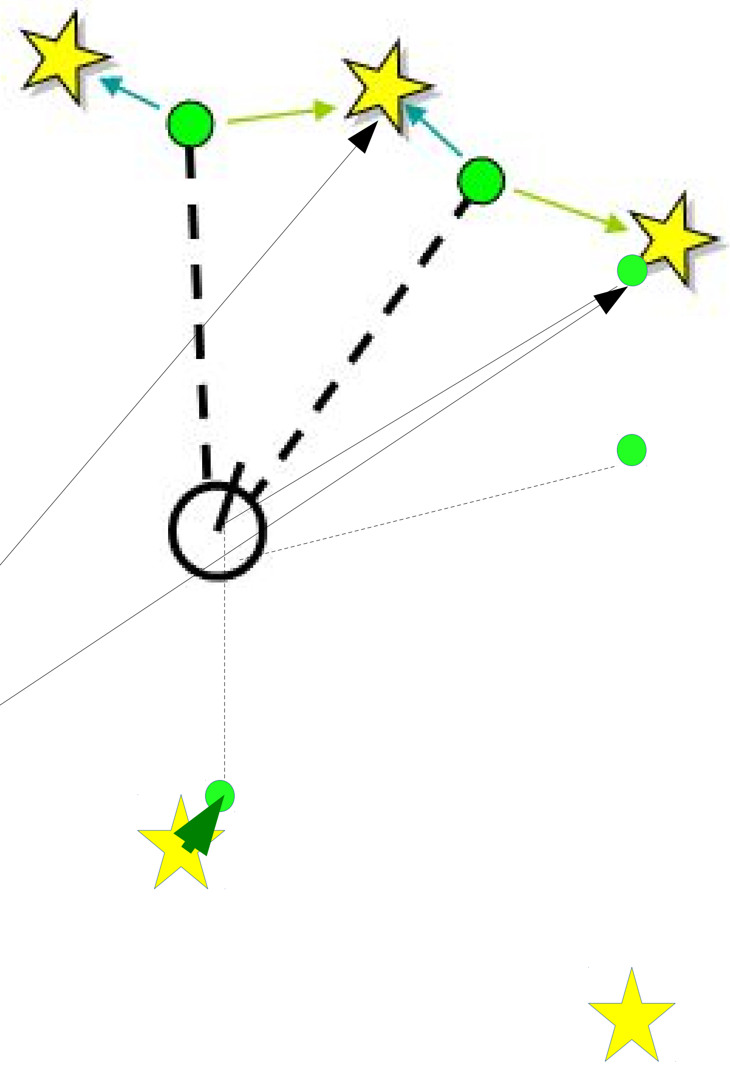


Pruning Heuristics

Bad associations are EVIL

To avoid the above cases we can use three heuristics

- Gating: ignore all associations whose cost is higher than a threshold
- Best friends: an association should be the best of both row and column
- Lonely best friends: the measurement should be clearly assigned to a single landmark and viceversa.



Implementing Heuristics

They are easier to implement

- **Gating**: ignore all associations with a cost higher than a certain value
- **Best friends**: an association should be the minimum of the row **and** of the column
- **Lonely best friends**: the difference between the minimum and the second minimum of a row/column should be higher than a threshold

All associations that fail the test are dropped and the measurements ignored

Summary

Data association is hard because the measurements are not independent without knowing the state

It would require to evaluate all possible assignments

Still errors are possible

Forcing independence renders the problem tractable in polynomial time

Acceptable solutions for SLAM can be found through greedy assignments, and potentially ambiguous situations can be avoided by dropping measurements