Robotic Motion Planning: Controls Primer

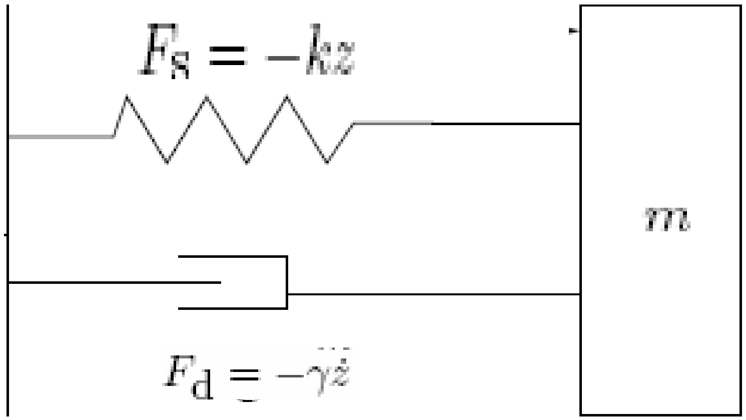
Robotics Institute 16-735
http://voronoi.sbp.ri.cmu.edu/~motion

Howie Choset
http://voronoi.sbp.ri.cmu.edu/~choset

State Space

$$m\ddot{z}(t) = -\gamma\dot{z}(t) - kz(t)$$

Model of mass spring damper system



z(t) position, *ż(t)* velocity
t_0 initial time, $z(t_0)$, $\dot{z}(t_0)$ initial position & velocity

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} z(t) \\ \dot{z}(t) \end{bmatrix}$$

$$\dot{x}(t) = \begin{bmatrix} \dot{z}(t) \\ \ddot{z}(t) \end{bmatrix} = \begin{bmatrix} x_2(t) \\ -\frac{1}{m}\left(\gamma\dot{z}(t) + kz(t)\right) \end{bmatrix} = \begin{bmatrix} x_2(t) \\ -\frac{1}{m}\left(\gamma x_2(t) + kx_1(t)\right) \end{bmatrix}$$

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{\gamma}{m} \end{bmatrix} x(t)$$

State space representation of mass spring damper system (1st order ODE)

$$\dot{x}(t) = Ax(t); \qquad x(t_0) = x_0$$

$$x(t) \in \mathbb{R}^n \text{ and } A \in \mathbb{R}^{n\times n}$$

Vector Field

assigns a vector Ax to each point x in the state space.

$$x(t) = e^{A(t-t_0)}x_0$$

$$\begin{aligned} e^{A(t-t_0)} &= \sum_{i=0}^{\infty} \frac{A^i (t-t_0)^i}{i!} \\ &= I_{n\times n} + A(t-t_0) + \frac{A^2 (t-t_0)^2}{2!} + \cdots \end{aligned}$$

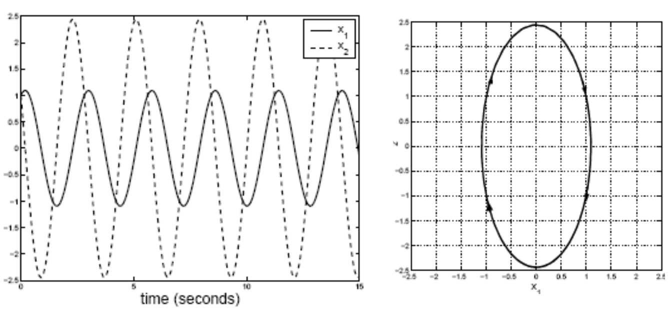
Stability

$$x(t) = e^{A(t-t_0)}x_0 \qquad \dot{x}(t) = Ax(t); \qquad x(t_0) = x_0$$

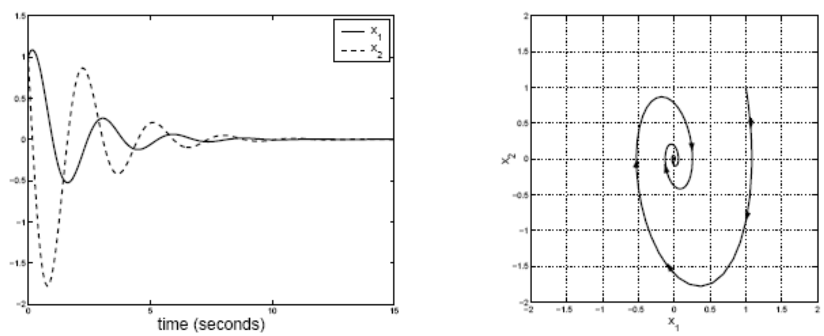
Equilibrium $\dot{x} = 0$ which occurs here at $x_e = 0$

for every $\varepsilon > 0$, there exists a δ for initial conditions $\|x_e - x(t_0)\| < \delta$

Stability:
$x(t)$ satisfies $\|x_e - x(t)\| < \epsilon$



Asymptotic Stability:
$\|x_e - x(t)\| \to 0$ as $t \to \infty$



Unstable: Neither

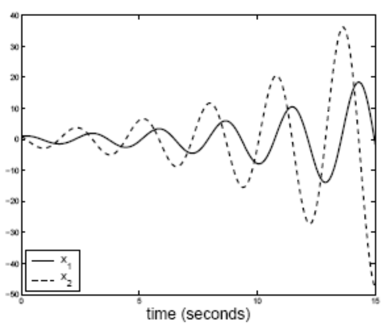
Stability and Eigenvalues

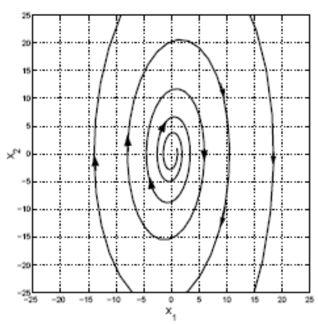
let λ_i, $i \in \{1, 2, \ldots, n\}$ denote the eigenvalues of A. Let $\mathrm{re}(\lambda_i)$ denote the real part of λ_i Then the following holds:

1. $x_e = 0$ is stable if and only if $\mathrm{re}(\lambda_i) \leq 0$ for all i.
2. $x_e = 0$ is asymptotically stable if and only if $\mathrm{re}(\lambda_i) < 0$ for all i.
3. $x_e = 0$ is unstable if and only if $\mathrm{re}(\lambda_i) > 0$ for any i.

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{\gamma}{m} \end{bmatrix} x(t)$$

$$\frac{-\gamma \pm \sqrt{\gamma^2 - 4km}}{2m}.$$





Negative damping….

Apply a force (a control input)

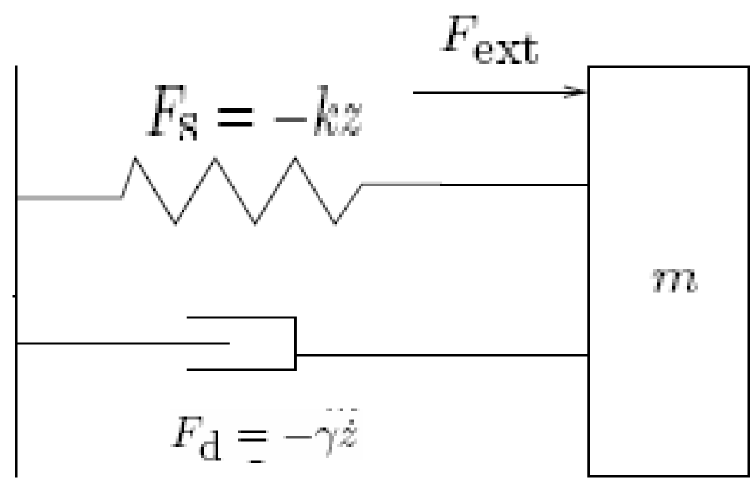
$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{\gamma}{m} \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} F(t)$$

$$\dot{x}(t) = Ax(t) + Bu(t); \qquad x(t_0) = x_0.$$

$x(t) \in \mathbb{R}^n$

$u(t) \in \mathbb{R}^m$

$B \in \mathbb{R}^{n \times m}$



Controlability

For any initial condition $x(t_0)$, there exists an input $u(t)$ that drives the solution $x(t)$ to the origin*

$$\dot{x}(t) = Ax(t) + Bu(t); \qquad x(t_0) = x_0$$

The LTI is control system controllable if and only if W has rank n where

$$W = [\, B \; AB \; A^2B \; \dots \; A^{n-1}B \,]$$

*(assuming the origin is an equilibrium point for the unforced system)

Closed Loop Stability: Pole Placement

Make an unforced unstable system stable

$$\dot{x}(t) = Ax(t) + Bu(t)$$

State dependent control law

$$u(t) = -Kx(t) \quad K \in \mathbb{R}^{m \times n}$$

$$\dot{x}(t) = \boxed{(A - BK)}\, x(t)$$

look at the eigenvalues

For a real valued matrix, if an eigenvalue *a* + *ib* has *b* ≠ *0*, then *a* – *ib* is an eigenvalue

So,

$$\bar{\Lambda} = \{\lambda_i | i \in \{1, 2, \ldots, n\}\}$$

Is ***allowable*** if for each λ_i that has an imaginary part, there is a λ_j that is a complex conjugate

<u>Assume</u>

This system is controllable, i.e., the pair (A, B) is controllable

B is full rank

$\bar{\Lambda} = \{\bar{\lambda}_i | i \in \{1, 2, \ldots, n\}\}$ is an allowable set of complex numbers

Then there exists an constant matrix $K \in \mathbb{R}^{m \times n}$ so that $(A - BK)$ *is equal to* Λ

What does this mean?

- To stabilize

$$\dot{x}(t) = Ax(t) + Bu(t)$$

State dependent control law

$$u(t) = -Kx(t) \quad K \in \mathbb{R}^{m \times n}$$

$$\dot{x}(t) = (A - BK)x(t)$$

- Find a *K* so that the Eigenvalues of *A – BK* have negative real values and are complex conjugates.

- The famous LQR does this by optimizing a user-defined cost function.

-

Example

- Mass-spring-damper (negative damping - why)

$$A - BK = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{\gamma}{m} \end{bmatrix} - \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} [k_1 \ k_2] = \begin{bmatrix} 0 & 1 \\ -\frac{k+k_1}{m} & \frac{\gamma+k_2}{m} \end{bmatrix}$$

- Eigenvalues

$$\frac{(-\gamma - k_2) \pm \sqrt{(-\gamma - k_2)^2 - 4(k - k_1)m}}{2m},$$

- Chose $k_2 \text{ such that } -\gamma - k_2 < 0$

- This is like adding positive damping

Observing LTI Systems

- Sadly, one cannot always sense all of the state variables

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t); \qquad x(t_0) = x_0, \\ y(t) &= Cx(t), \end{aligned}$$

state	*control*	*output*	
$x(t) \in \mathbb{R}^n$	$u(t) \in \mathbb{R}^m$	$y(t) \in \mathbb{R}^p$	$C \in \mathbb{R}^{p \times n}$.

May not be invertible (nor square)

- Example: can only sense or measure velocity

$$\begin{aligned} \dot{x}(t) &= \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{\gamma}{m} \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} F(t), \\ y(t) &= [0 \; 1] \, x(t), \end{aligned}$$

- Can we recover the state from the observations??

Observable

*A system is **observable** if one can determine the initial state by observing the output and knowing the controls over some period of time*

The system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t); \qquad x(t_0) = x_0 \\ y(t) &= Cx(t), \end{aligned}$$

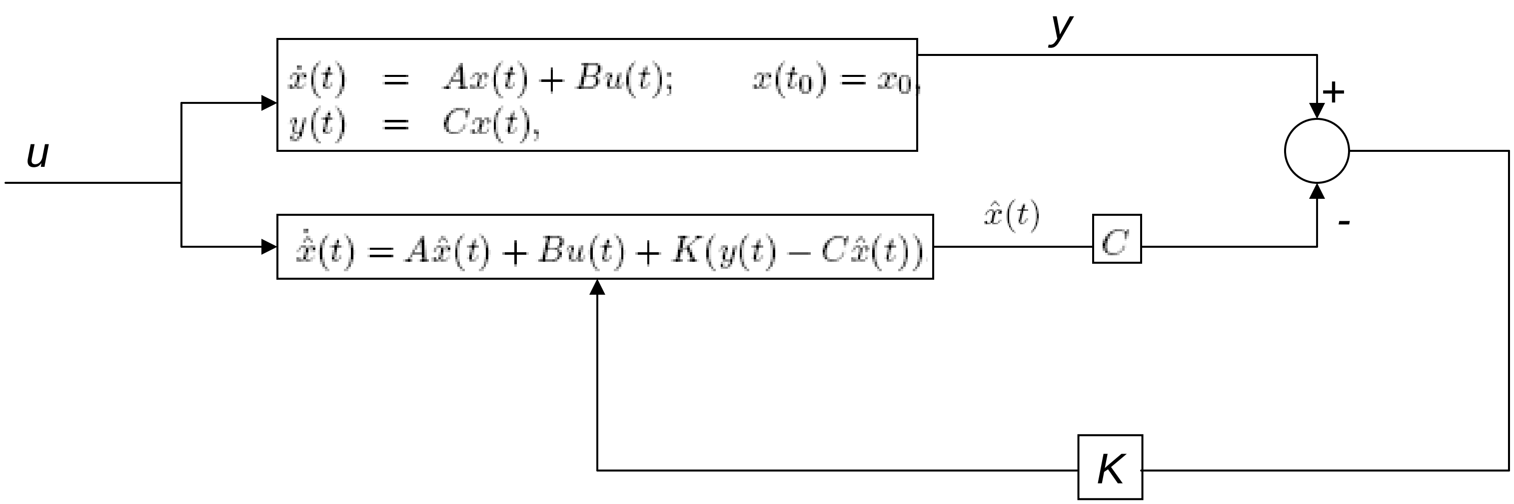
Is observable if $\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$ has rank n

The pair (A,C) is considered observable.

(A,C) observable if and only if (A^T, C^T) is controllable

If (A,B) observable and (A,C) observable, then (A,B,C) is ***minimal***

Observer



$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + K(y(t) - C\hat{x}(t))$$

We know

State estimate

$$\hat{x}(t)$$

Estimated output

$$C\hat{x}(t)$$

Copy of original system with a correcting term $K(y(t) - C\hat{x}(t))$
which is the difference between the output and estimated output

How does this work

- Consider the error $e(t) = x(t) - \hat{x}(t)$

$$\begin{aligned}
\dot{e}(t) &= \dot{x}(t) - \dot{\hat{x}}(t) \\
&= Ax(t) + Bu(t) - (A\hat{x}(t) + Bu(t) + K(y(t) - C\hat{x}(t))) \\
&= A(x(t) - \hat{x}(t)) - K(Cx(t) - C\hat{x}(t)) \\
&= (A - KC)e(t)
\end{aligned}$$

- As $e(t) \to 0, \text{ then } \hat{x}(t) \to x(t)$
- Chose *K* so that $\dot{e}(t) = (A - KC)e(t)$ asymp. Stable
- Look at eigenvalues of $A - KC$ (not quite right form)
- Look at eigenvalues of $A^t - C^t K^t$ (same eigenvalues)
- Make sure such eigenvalues are allowable
  - Make sure (A^t, C^t) is controllable
  - C^t is full rank
- Same as saying (A,C) is observable
- So it is a matter of choosing *K*

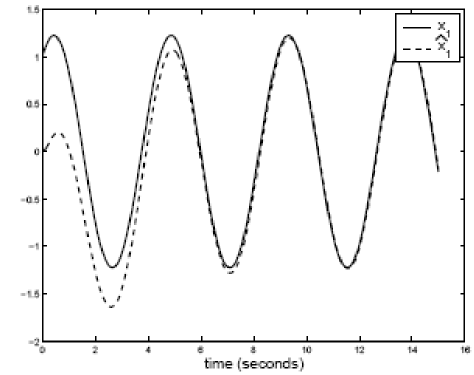
Example

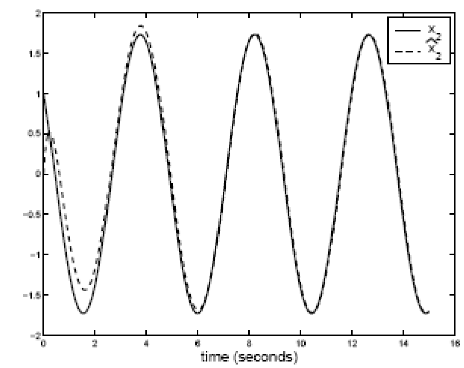
$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{\gamma}{m} \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} F(t),$$
$$y(t) = [0\ 1]\, x(t),$$

$$A - KC = \begin{bmatrix} 0 & 1 - k_1 \\ -\frac{k}{m} & -\frac{\gamma}{m} - k_2 \end{bmatrix}$$

$$\frac{-(\gamma + mk_2) \pm \sqrt{(-\gamma - mk_2)^2 - 4m(k - k_1)}}{2m}$$

Chose k_2 such that $-\gamma - mk_2 < 0$





$m = 1,\ k = 2,$ and $\gamma = 0,$

Chose $K = [0\ 2]^T$

Discrete Time

$$\dot{x}(t_0 + kT) \approx \frac{x(k+1) - x(k)}{T} \approx Ax(k) + Bu(k)$$

$$x(k+1) \approx x(k) + TAx(k) + TBu(k)$$

$F = I_{n \times n} + TA$, $G = TB$, and $H = C$

$$\begin{aligned} \dot{x}(k+1) &= Fx(k) + Gu(k); \qquad x(0) = x_0 \\ y(k) &= Hx(k) \end{aligned}$$

Properties

- Stability of *x(k+1) = F(x) x(k)*
  - *Let* $\lambda_i,\ i \in \{1, 2, \ldots, n\}$ *denote the eigenvalues of F*

$x_e = 0$ *is stable if and only if* $|\lambda_i| \leq 1$ *for all* i.

$x_e = 0$ *is asymptotically stable if and only if* $|\lambda_i| < 1$ *for all* i

$x_e = 0$ *is unstable if and only if* $|\lambda_i| > 1$ *for any* i.

- Controllability of *(F,G)*

$$[G\ FG\ F^2G\ \cdots\ F^{n-1}G] \text{ has rank } n$$

- *Observability of (F,H) – if* F^t, H^t *is controllable*

Localization, Mapping, SLAM and The Kalman Filter according to George

Robotics Institute 16-735

http://voronoi.sbp.ri.cmu.edu/~motion

Howie Choset

http://voronoi.sbp.ri.cmu.edu/~choset

The Problem

- What is the world around me (mapping)
  - sense from various positions
  - integrate measurements to produce map
  - assumes perfect knowledge of position
- Where am I in the world (localization)
  - sense
  - relate sensor readings to a world model
  - compute location relative to model
  - assumes a perfect world model
- Together, these are SLAM (Simultaneous Localization and Mapping)

Localization

Tracking: Known initial position

Global Localization: Unknown initial position

Re-Localization: Incorrect known position

(kidnapped robot problem)

SLAM

Mapping while tracking locally and globally

Challenges

- Sensor processing
- Position estimation
- Control Scheme
- Exploration Scheme
- Cycle Closure
- Autonomy
- Tractability
- Scalability

Representations for Robot Localization

<u>Discrete approaches ('95)</u>
- Topological representation ('95)
  - uncertainty handling (POMDPs)
  - occas. global localization, recovery
- Grid-based, metric representation ('96)
  - global localization, recovery

<u>Kalman filters (late-80s?)</u>
- Gaussians
- approximately linear models
- position tracking

Robotics

AI

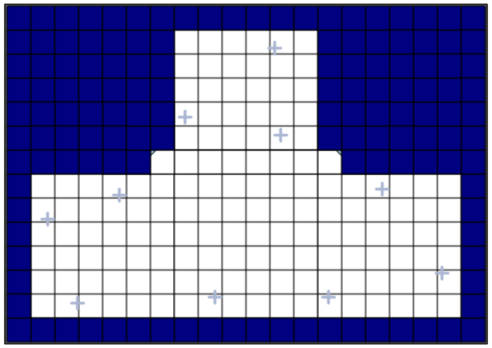
<u>Particle filters ('99)</u>
- sample-based representation
- global localization, recovery

<u>Multi-hypothesis ('00)</u>
- multiple Kalman filters
- global localization, recovery

Three Major Map Models

Grid-Based:

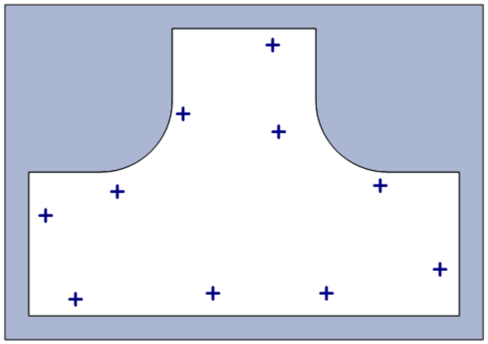
Collection of discretized obstacle/free-space pixels



Elfes, Moravec, Thrun, Burgard, Fox, Simmons, Koenig, Konolige, etc.

Feature-Based:

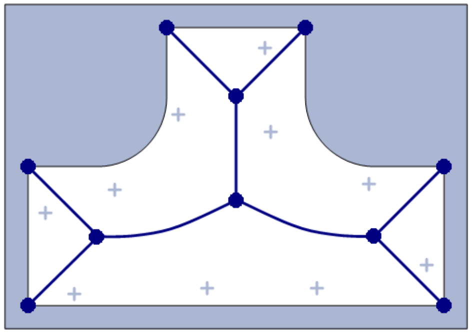
Collection of landmark locations and correlated uncertainty



Smith/Self/Cheeseman, Durrant–Whyte, Leonard, Nebot, Christensen, etc.

Topological:

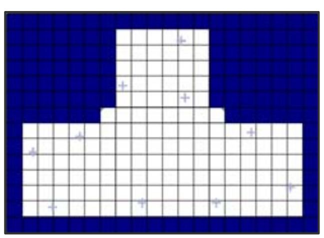
Collection of nodes and their interconnections

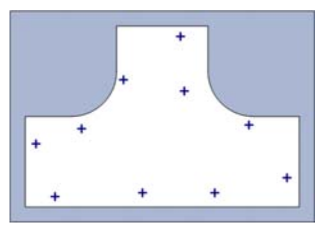


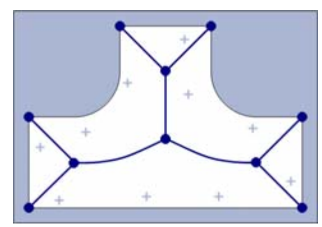
Kuipers/Byun, Chong/Kleeman, Dudek, Choset, Howard, Mataric, etc.

Three Major Map Models

	Grid-Based	Feature-Based	Topological
Resolution vs. Scale	Discrete localization	Arbitrary localization	Localize to nodes
Computational Complexity	Grid size and resolution	Landmark covariance (N^2)	Minimal complexity
Exploration Strategies	Frontier-based exploration	No inherent exploration	Graph exploration



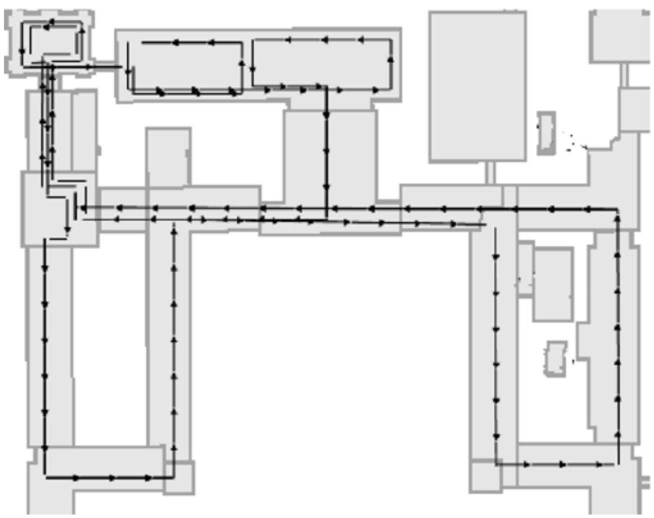


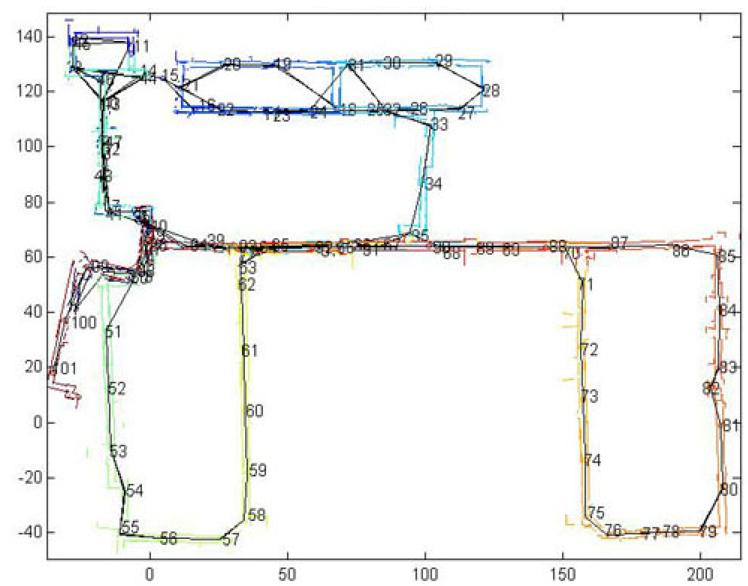


Atlas Framework

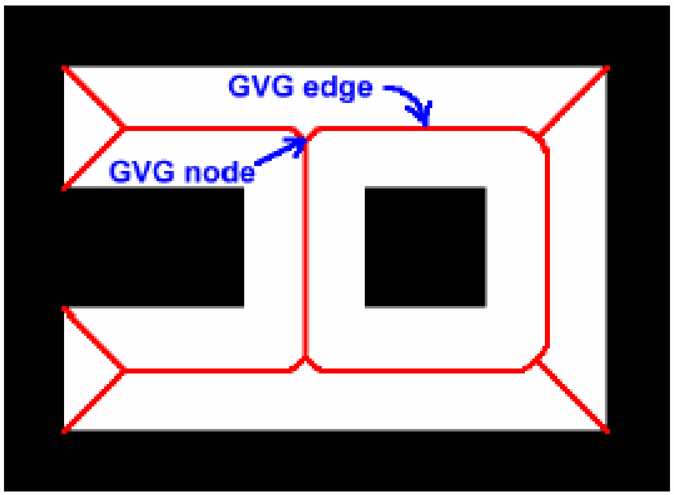
- **Hybrid Solution:**
  - Local features extracted from local grid map.
  - Local map frames created at complexity limit.
  - Topology consists of connected local map frames.

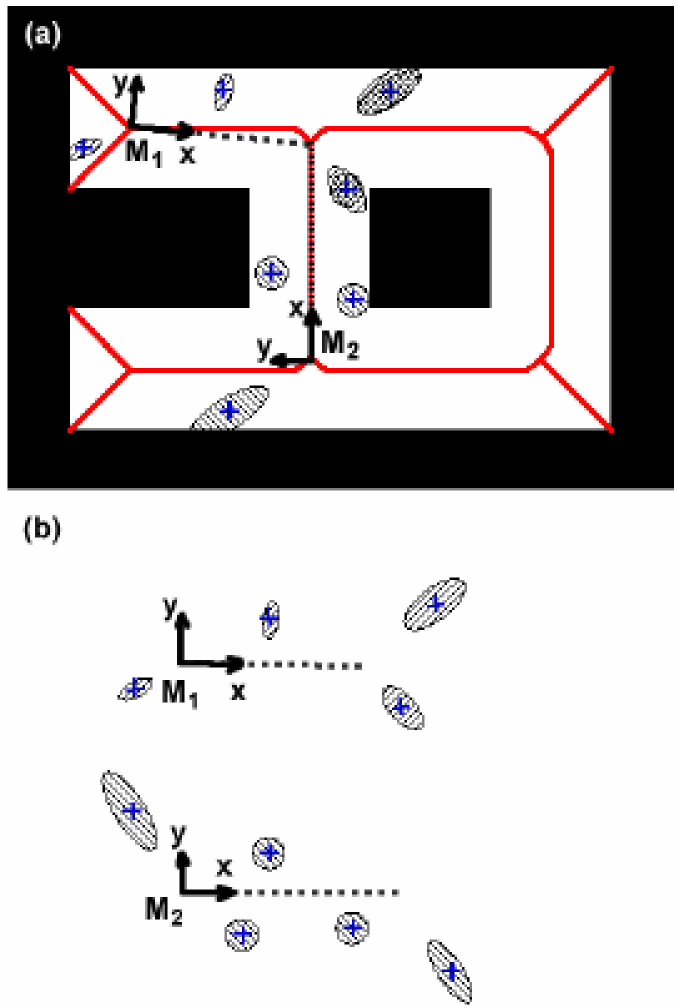
Authors: Chong, Kleeman; Bosse, Newman, Leonard, Soika, Feiten, Teller





H-SLAM





What does a Kalman Filter do, anyway?

Given the linear dynamical system:

$$x(k+1) = F(k)x(k) + G(k)u(k) + v(k)$$

$$y(k) = H(k)x(k) + w(k)$$

$x(k)$ is the n-dimensional state vector (unknown)

$u(k)$ is the m-dimensional input vector (known)

$y(k)$ is the p-dimensional output vector (known, measured)

$F(k), G(k), H(k)$ are appropriately dimensioned system matrices (known)

$v(k), w(k)$ are zero-mean, white Gaussian noise with (known) covariance matrices $Q(k), R(k)$

the Kalman Filter is a recursion that provides the "best" estimate of the state vector *x*.

What's so great about that?

$$x(k+1) = F(k)x(k) + G(k)u(k) + v(k)$$
$$y(k) = H(k)x(k) + w(k)$$

- noise smoothing (improve noisy measurements)
- state estimation (for state feedback)
- recursive (computes next estimate using only most recent measurement)

How does it work?

$$x(k+1) = F(k)x(k) + G(k)u(k) + v(k)$$
$$y(k) = H(k)x(k) + w(k)$$

1. prediction based on last estimate:

$$\hat{x}(k+1 \mid k) = F(k)\hat{x}(k \mid k) + G(k)u(k)$$
$$\hat{y}(k) = H(k)\hat{x}(k+1 \mid k)$$

2. calculate correction based on prediction and current measurement:

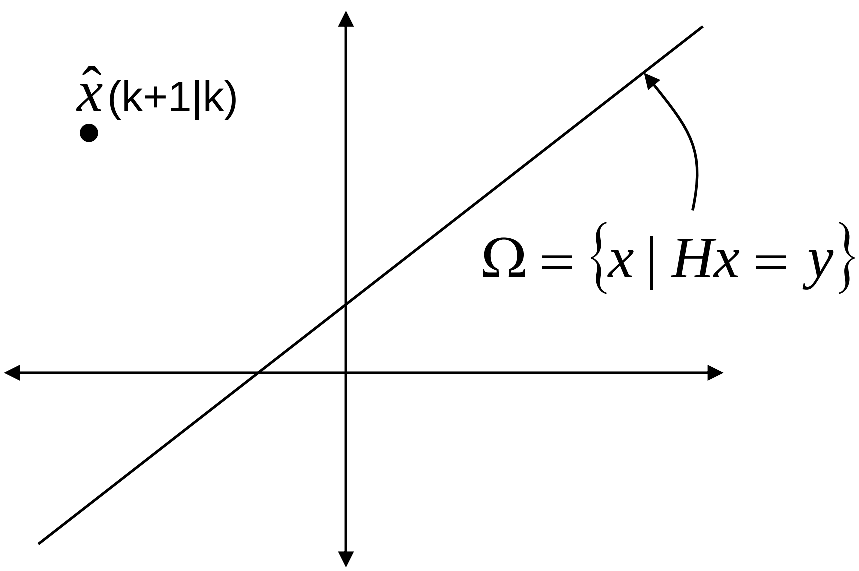
$$\Delta x = f\left(y(k+1), \hat{x}(k+1 \mid k)\right)$$

3. update prediction: $\hat{x}(k+1 \mid k+1) = \hat{x}(k+1 \mid k) + \Delta x$

Finding the correction (no noise!)

$$y = Hx$$

Given prediction $\hat{x}(k+1|k)$ and output y, find Δx so that $\hat{x} = \hat{x}(k+1|k) + \Delta x$ is the "best" estimate of x.



Want the best estimate to be consistent with sensor readings

"best" estimate comes from shortest Δx

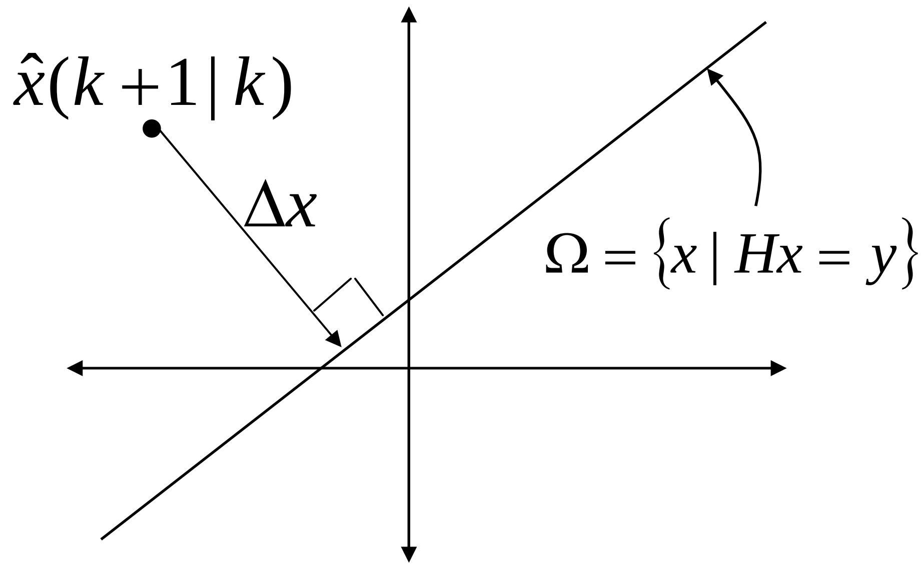
Finding the correction (no noise!)

$$y = Hx$$

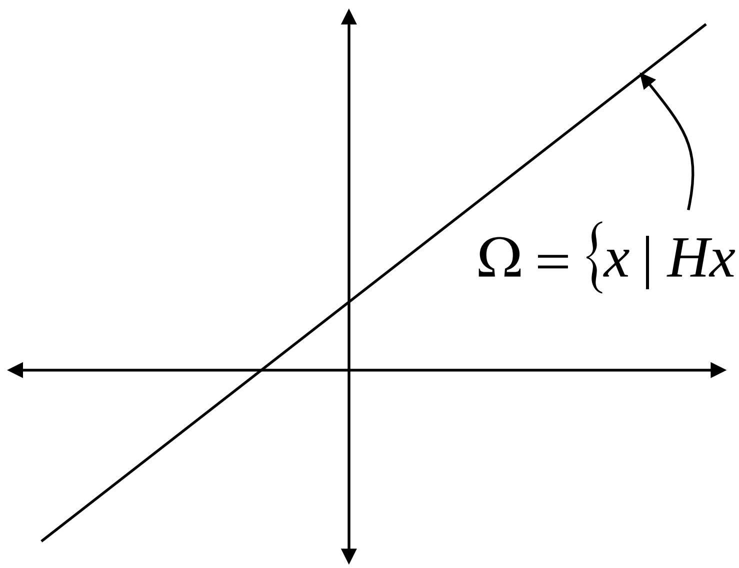
Given prediction $\hat{x}(k+1|1)$ and output y, find Δx so that $\hat{x} = \hat{x}(k+1|1) + \Delta x$ is the "best" estimate of x.

"best" estimate comes from shortest Δx

shortest Δx is perpendicular to Ω



Some linear algebra



$$\Omega = \{x \mid Hx = y\}$$

a is parallel to Ω if Ha = 0

$$Null(H) = \{a \neq 0 \mid Ha = 0\}$$

a is parallel to Ω if it lies in the null space of H

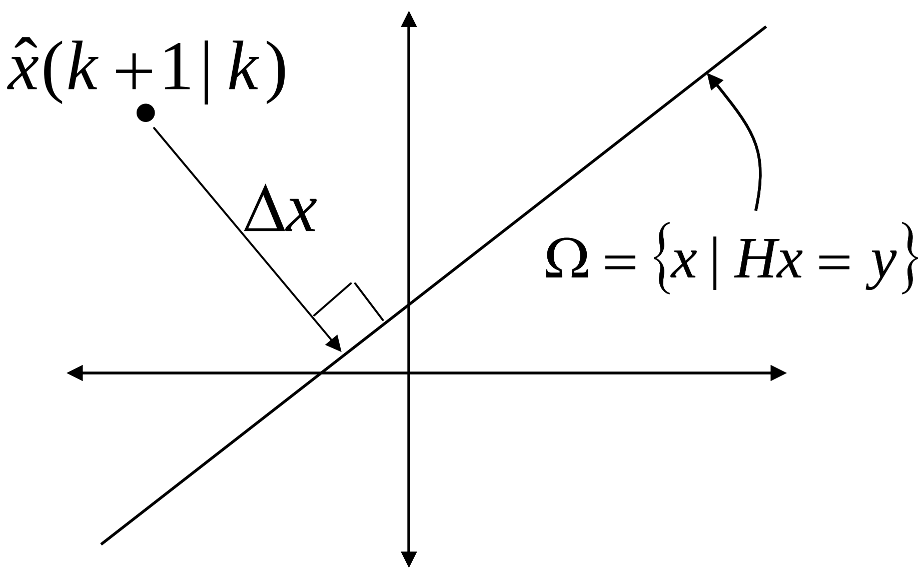
$$\text{for all } v \in Null(H), v \perp b \text{ if } b \in column(H^T)$$

Weighted sum of columns means $b = H\gamma$, the weighted sum of columns

Finding the correction (no noise!)

$$y = Hx$$

Given prediction $\hat{x}(k+1|k)$ and output y, find Δx so that $\hat{x} = \hat{x}(k+1|k) + \Delta x$ is the "best" estimate of x.



"best" estimate comes from shortest Δx

shortest Δx is perpendicular to Ω

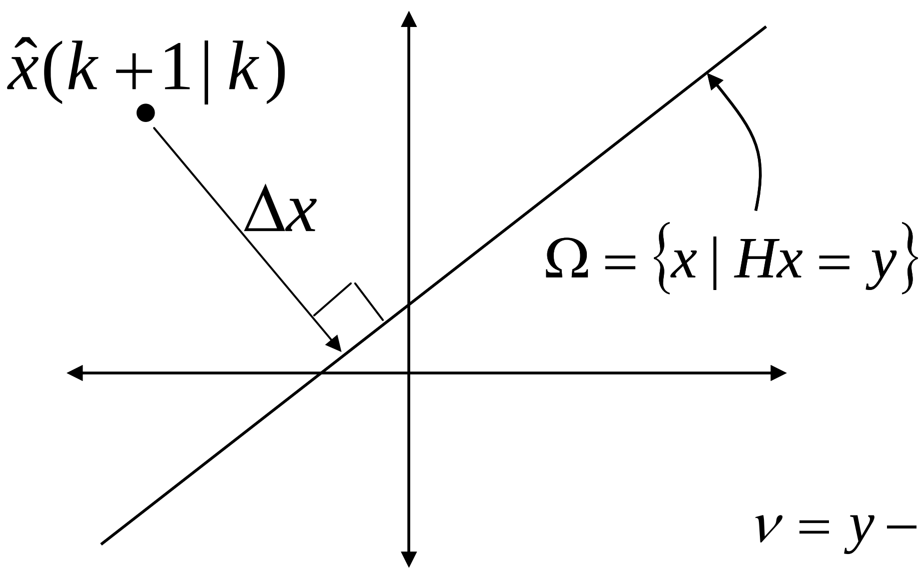
$$\Rightarrow \Delta x \in \text{null}(H)^{\perp} \Rightarrow \Delta x \in \text{column}(H^T)$$

$$\Rightarrow \Delta x = H^T \gamma$$

Finding the correction (no noise!)

$$y = Hx$$

Given prediction $\hat{x}(k+1|k)$ and output y, find Δx so that $\hat{x} = \hat{x}(k+1|k) + \Delta x$ is the "best" estimate of x.



"best" estimate comes from shortest Δx

shortest Δx is perpendicular to Ω

$$\Rightarrow \Delta x \in \text{null}(H)^{\perp} \Rightarrow \Delta x \in \text{column}(H^T)$$

$$\Rightarrow \Delta x = H^T \gamma$$

Real output – estimated output

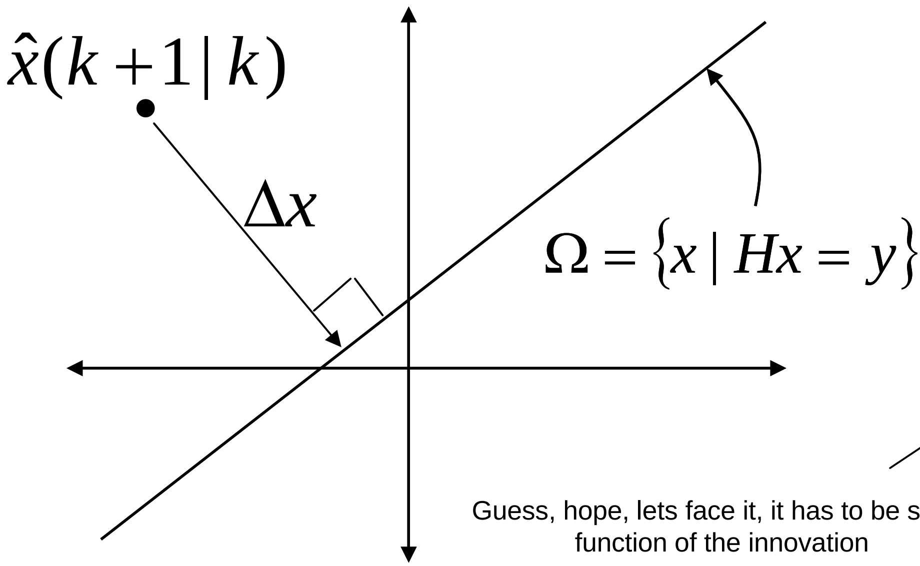
$$\nu = y - H(\hat{x}(k+1|k)) = H(x - \hat{x}(k+1|k))$$

innovation

Finding the correction (no noise!)

$$y = Hx$$

Given prediction $\hat{x}(k+1|k)$ and output y, find Δx so that $\hat{x} = \hat{x}(k+1|k) + \Delta x$ is the "best" estimate of x.



"best" estimate comes from shortest Δx

shortest Δx is perpendicular to Ω

$$\Rightarrow \Delta x \in \text{null}(H)^{\perp} \Rightarrow \Delta x \in \text{column}(H^T)$$

$$\Rightarrow \Delta x = H^T \gamma$$

assume γ is a linear function of ν

Guess, hope, lets face it, it has to be some function of the innovation

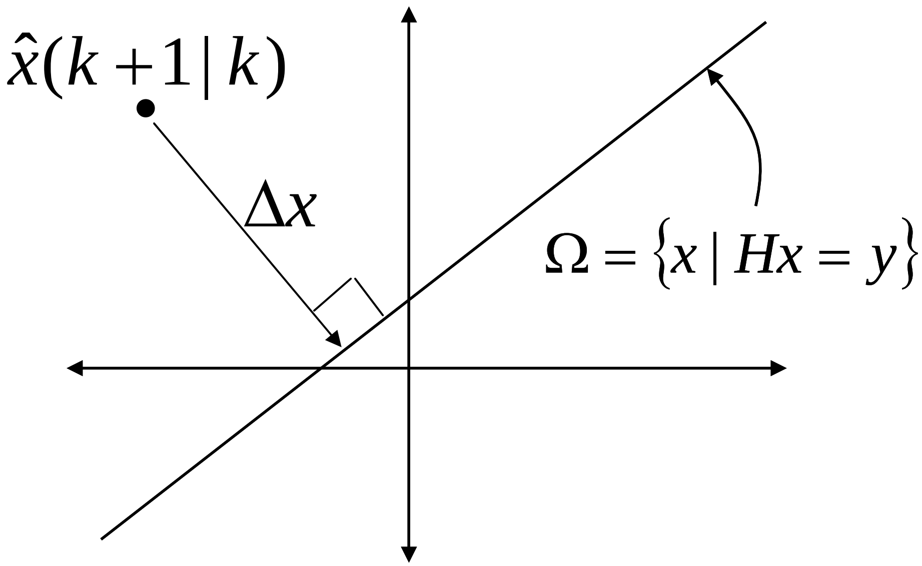
$$\Rightarrow \Delta x = H^T K \nu$$

for some $m \times m$ matrix K

Finding the correction (no noise!)

$$y = Hx$$

Given prediction $\hat{x}_-$ and output y, find Δx so that $\hat{x} = \hat{x}(k+1|k) + \Delta x$ is the "best" estimate of x.



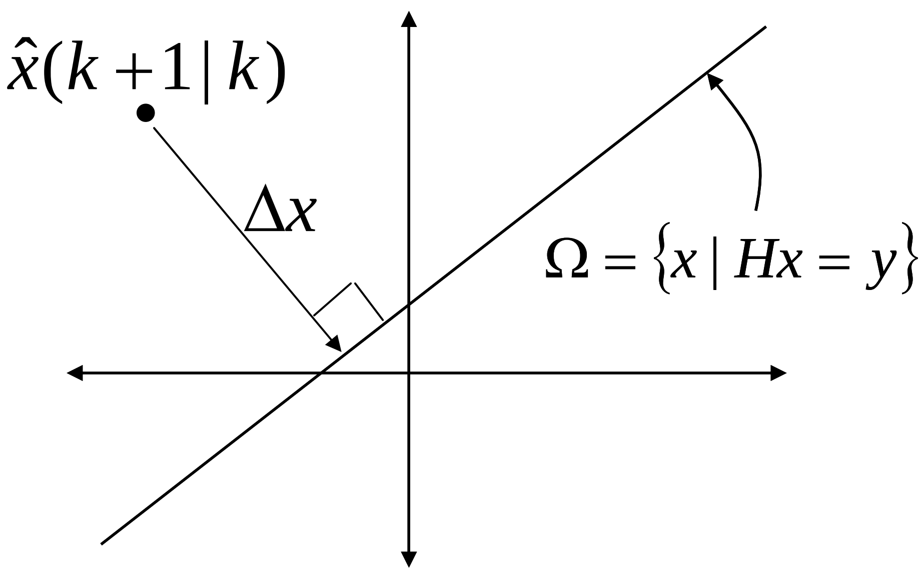
we require

$$H(\hat{x}(k+1|k) + \Delta x) = y$$

Finding the correction (no noise!)

$$y = Hx$$

Given prediction $\hat{x}_-$ and output y, find Δx so that $\hat{x} = \hat{x}(k+1 \mid k) + \Delta x$ is the "best" estimate of x.



we require

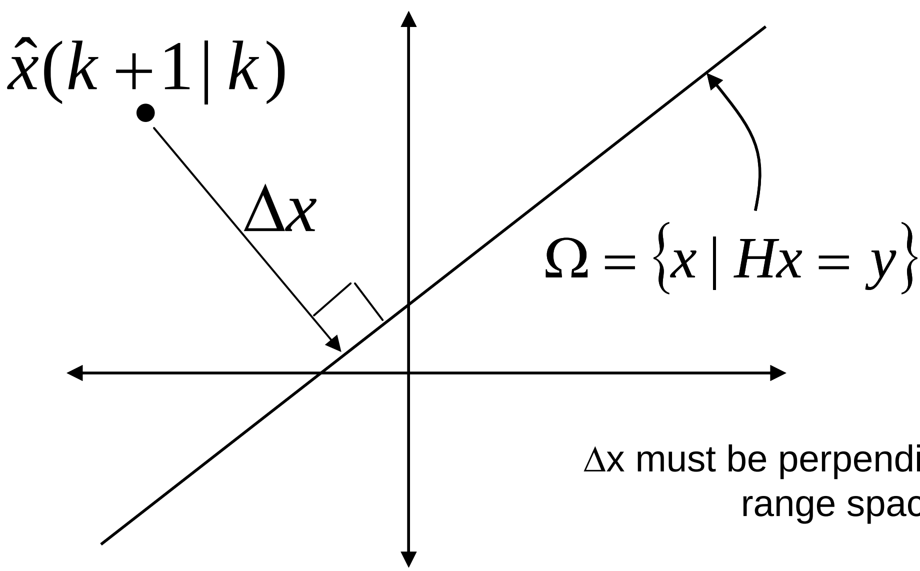
$$H(\hat{x}(k+1 \mid k) + \Delta x) = y$$

$$\Rightarrow H\Delta x = y - H\hat{x}(k+1 \mid k) = H(x - \hat{x}(k+1 \mid k)) = \nu$$

Finding the correction (no noise!)

$$y = Hx$$

Given prediction $\hat{x}_-$ and output y, find Δx so that $\hat{x} = \hat{x}(k+1|k) + \Delta x$ is the "best" estimate of x.



we require

$$H(\hat{x}(k+1|k) + \Delta x) = y$$

$$\Rightarrow H\Delta x = y - H\hat{x}(k+1|k) = H(x - \hat{x}(k+1|k)) = \nu$$

substituting $\Delta x = H^T K \nu$ yields

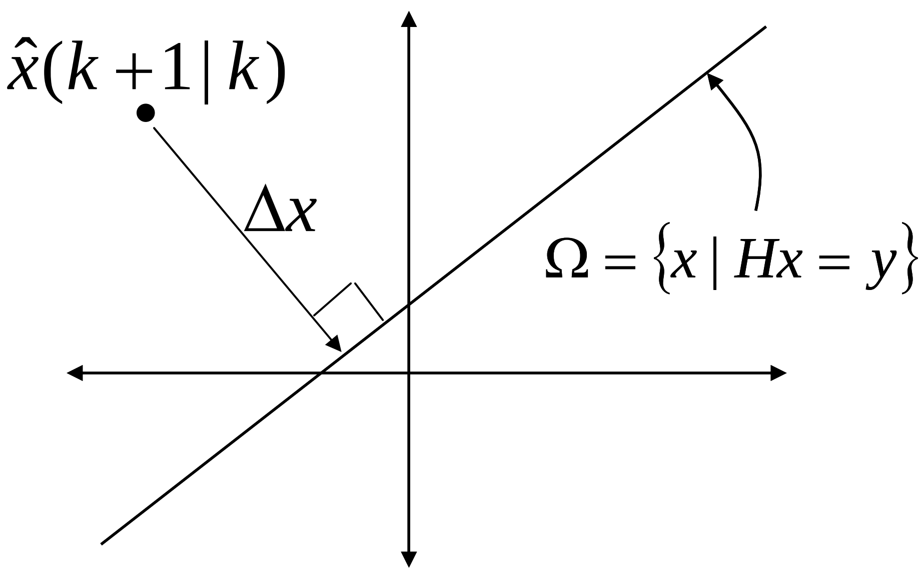
$$HH^T K \nu = \nu$$

Δx must be perpendicular to Ω b/c anything in the range space of H^T is perp to Ω

Finding the correction (no noise!)

$$y = Hx$$

Given prediction $\hat{x}_-$ and output y, find Δx so that $\hat{x} = \hat{x}(k+1|k) + \Delta x$ is the "best" estimate of x.



we require

$$H(\hat{x}(k+1|k) + \Delta x) = y$$

$$\Rightarrow H\Delta x = y - H\hat{x}(k+1|k) = H(x - \hat{x}(k+1|k)) = \nu$$

substituting $\Delta x = H^T K \nu$ yields

$$HH^T K \nu = \nu$$

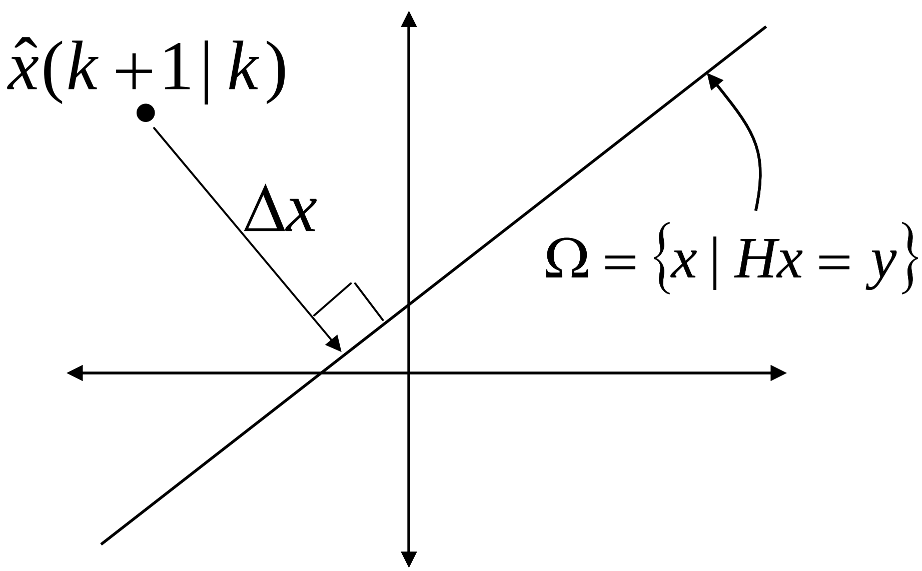
$$\Rightarrow K = \left(HH^T\right)^{-1}$$

The fact that the linear solution solves the equation makes assuming K is linear a kosher guess

Finding the correction (no noise!)

$$y = Hx$$

Given prediction $\hat{x}_-$ and output y, find Δx so that $\hat{x} = \hat{x}(k+1|k) + \Delta x$ is the "best" estimate of x.



we require

$$H(\hat{x}(k+1|k) + \Delta x) = y$$

$$\Rightarrow H\Delta x = y - H\hat{x}(k+1|k) = H(x - \hat{x}(k+1|k)) = \nu$$

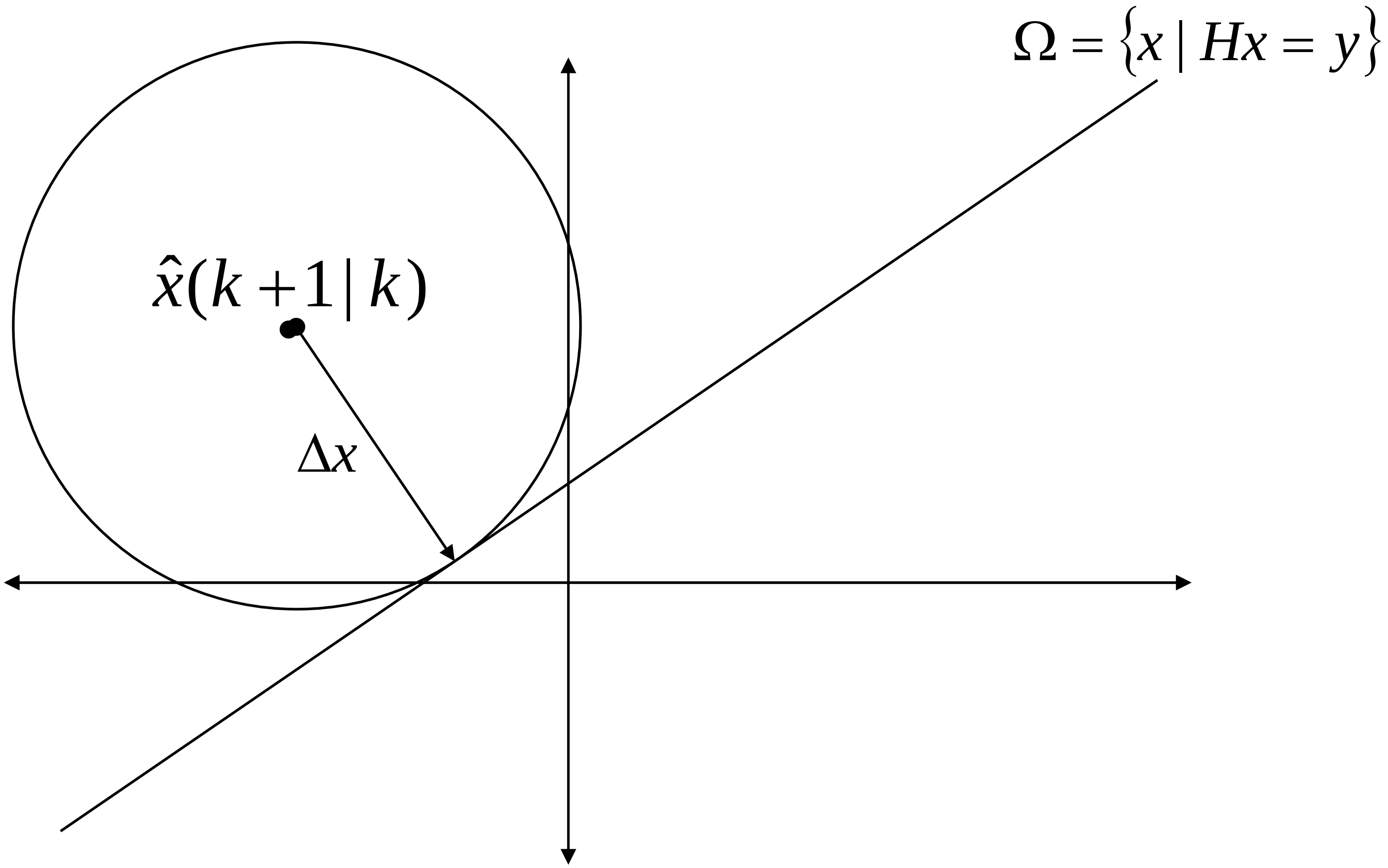
substituting $\Delta x = H^T K \nu$ yields

$$HH^T K \nu = \nu$$

$$\Rightarrow K = \left(HH^T\right)^{-1}$$

$$\Delta x = H^T \left(HH^T\right)^{-1} \nu$$

A Geometric Interpretation



A Simple State Observer

System:

$$x(k+1) = Fx(k) + Gu(k)$$

$$y(k) = Hx(k)$$

Observer:

1. prediction:

$$\hat{x}(k+1 \mid k) = F\hat{x}(k \mid k) + Gu(k)$$

2. compute correction:

$$\Delta x = H^T \left(HH^T\right)^{-1} \left(y(k+1) - H\hat{x}(k+1 \mid k)\right)$$

3. update:

$$\hat{x}(k+1 \mid k+1) = \hat{x}(k+1 \mid k) + \Delta x$$

Caveat #1

Note: The observer presented here is not a very good observer. Specifically, it is not guaranteed to converge for all systems. Still the intuition behind this observer is the same as the intuition behind the Kalman filter, and the problems will be fixed in the following slides.

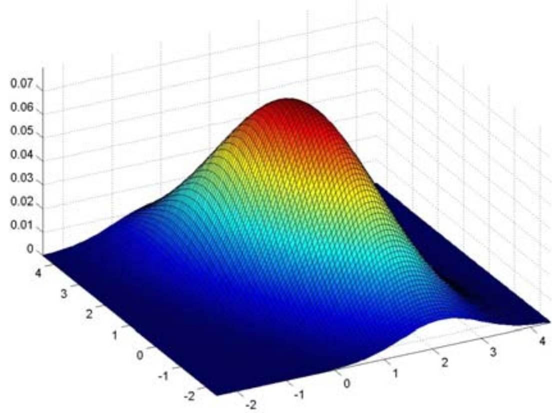
It really corrects only to the current sensor information, so if you are on the hyperplane but not at right place, you have no correction…. **I am waiving my hands here, look in book**

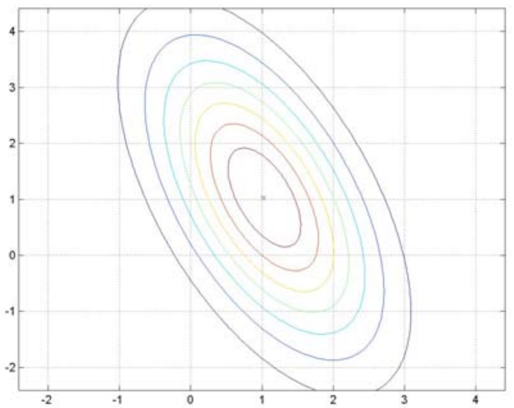
Estimating a *distribution* for *x*

Our estimate of *x* is not exact!
We can do better by estimating a joint Gaussian distribution *p(x)*.

$$p(x) = \frac{1}{(2\pi)^{n/2}\left|P\right|^{1/2}} e^{\frac{-1}{2}\left((x-\hat{x})^T P^{-1} (x-\hat{x})\right)}$$

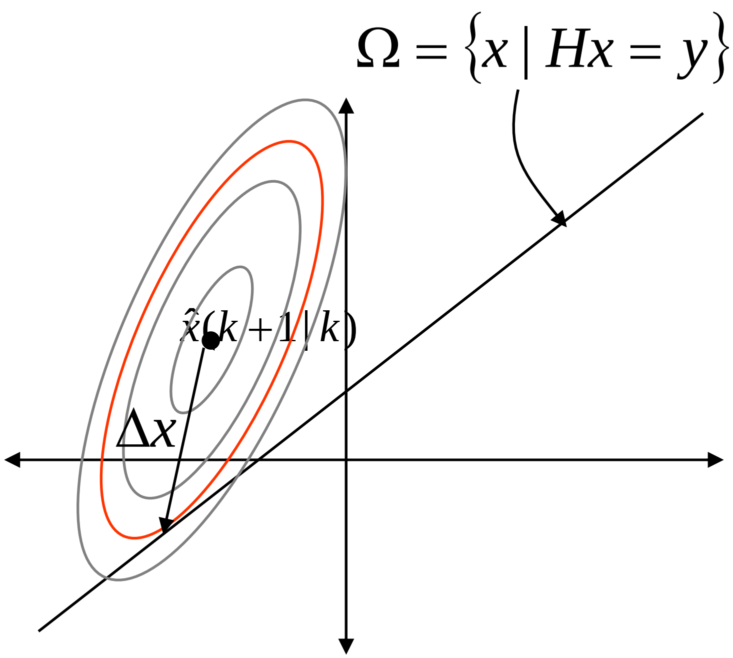
where $P = E\left((x-\hat{x})(x-\hat{x})^T\right)$ is the *covariance matrix*





Finding the correction (geometric intuition)

Given prediction $\hat{x}(k+1|k)$, covariance P, and output y, find Δx so that $\hat{x} = \hat{x}(k+1|k) + \Delta x$ is the "best" (i.e. most probable) estimate of x.



$$p(x) = \frac{1}{(2\pi)^{n/2}|P|^{1/2}} e^{\frac{-1}{2}\left((x-\hat{x})^T P^{-1}(x-\hat{x})\right)}$$

The most probable Δx is the one that :

1. satisfies $\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + \Delta x$
2. minimizes $\Delta x^T P^{-1} \Delta x$

A new kind of distance

Suppose we define a new inner product on R^n to be:

$$\langle x_1, x_2 \rangle = x_1^T P^{-1} x_2 \qquad \text{(this replaces the old inner product } x_1^T x_2\text{)}$$

Then we can define a new norm $\|x\|^2 = \langle x, x \rangle = x^T P^{-1} x$

The $\hat{x}$ in Ω that minimizes $\|\Delta x\|$ is the orthogonal projection of $\hat{x}(k+1 \mid k)$ onto Ω, so Δx is orthogonal to Ω.

$$\Rightarrow \langle \omega, \Delta x \rangle = 0 \text{ for } \omega \text{ in } \mathrm{T}\Omega = null(H)$$

$$\langle \omega, \Delta x \rangle = \omega^T P^{-1} \Delta x = 0 \text{ iff } \Delta x \in column(PH^T)$$

Finding the correction (for real this time!)

Assuming that Δx is linear in $\nu = y - H\hat{x}(k+1|k)$

$$\Delta x = PH^T K \nu$$

The condition $y = H(\hat{x}(k+1|k) + \Delta x) \Rightarrow H\Delta x = y - H\hat{x}(k+1|k) = \nu$

Substitution yields:

$$H\Delta x = \nu = HPH^T K \nu$$

$$\Rightarrow K = \left(HPH^T\right)^{-1}$$

$$\therefore \quad \Delta x = PH^T \left(HPH^T\right)^{-1} \nu$$

A Better State Observer

$$x(k+1) = Fx(k) + Gu(k) + v(k)$$

$$y(k) = Hx(k)$$

Sample of Guassian Dist. w/ COV Q

We can create a better state observer following the same 3. steps, but now we must also estimate the covariance matrix *P*.

We start with $x(k|k)$ and $P(k|k)$

Step 1: Prediction

Where did noise go? Expected value…

$$\hat{x}(k+1 \mid k) = F\hat{x}(k \mid k) + Gu(k)$$

What about *P*? From the definition:

$$P(k \mid k) = E\left((x(k) - \hat{x}(k \mid k))(x(k) - \hat{x}(k \mid k))^T\right)$$

and

$$P(k+1 \mid k) = E\left((x(k+1) - \hat{x}(k+1 \mid k))(x(k+1) - \hat{x}(k+1 \mid k))^T\right)$$

Continuing Step 1

To make life a little easier, lets shift notation slightly:

$$P^-_{k+1} = E\left((x_{k+1} - \hat{x}^-_{k+1})(x_{k+1} - \hat{x}^-_{k+1})^T\right)$$

$$= E\left(\left(Fx_k + Gu_k + v_k - (F\hat{x}_k + Gu_k)\right)\left(Fx_k + Gu_k + v_k - (F\hat{x}_k + Gu_k)\right)^T\right)$$

$$= E\left(\left(F\left(x_k - \hat{x}_k\right) + v_k\right)\left(F\left(x_k - \hat{x}_k\right) + v_k\right)^T\right)$$

$$= E\left(F\left(x_k - \hat{x}_k\right)\left(x_k - \hat{x}_k\right)^T F^T + 2F\left(x_k - \hat{x}_k\right)v_k^T + v_k v_k^T\right)$$

$$= FE\left(\left(x_k - \hat{x}_k\right)\left(x_k - \hat{x}_k\right)^T\right)F^T + E\left(v_k v_k^T\right)$$

$$= FP_k F^T + Q$$

$$P(k+1 \mid k) = FP(k \mid k)F^T + Q$$

Step 2: Computing the correction

From step 1 we get $\hat{x}(k+1|k)$ and $P(k+1|k)$.

Now we use these to compute Δx :

$$\Delta x = P(k+1|k)H\left(HP(k+1|k)H^T\right)^{-1}\left(y(k+1) - H\hat{x}(k+1|k)\right)$$

For ease of notation, define W so that

$$\Delta x = W\nu$$

Step 3: Update

$$\hat{x}(k+1 \mid k+1) = \hat{x}(k+1 \mid k) + W\nu$$

$$P_{k+1} = E\left((x_{k+1} - \hat{x}_{k+1})(x_{k+1} - \hat{x}_{k+1})^T\right)$$

$$= E\left((x_{k+1} - \hat{x}_{k+1}^{-} - W\nu)(x_{k+1} - \hat{x}_{k+1}^{-} - W\nu)^T\right)$$

(just take my word for it…)

$$P(k+1 \mid k+1) = P(k+1 \mid k) - WHP(k+1 \mid k)H^T W^T$$

Just take my word for it…

$$
\begin{aligned}
P_{k+1} &= E\left((x_{k+1}-\hat{x}_{k+1})(x_{k+1}-\hat{x}_{k+1})^T\right) \\
&= E\left((x_{k+1}-\hat{x}^-_{k+1}-W\nu)(x_{k+1}-\hat{x}^-_{k+1}-W\nu)^T\right) \\
&= E\left(\left((x_{k+1}-\hat{x}^-_{k+1})-W\nu\right)\left((x_{k+1}-\hat{x}^-_{k+1})-W\nu\right)^T\right) \\
&= E\left((x_{k+1}-\hat{x}^-_{k+1})(x_{k+1}-\hat{x}^-_{k+1})^T-2W\nu(x_{k+1}-\hat{x}^-_{k+1})^T+W\nu(W\nu)^T\right) \\
&= P^-_{k+1}+E\left(-2WH(x_{k+1}-\hat{x}^-_{k+1})(x_{k+1}-\hat{x}^-_{k+1})^T+WH(x_{k+1}-\hat{x}^-_{k+1})(x_{k+1}-\hat{x}^-_{k+1})^T H^T W^T\right) \\
&= P^-_{k+1}-2WHP^-_{k+1}+WHP^-_{k+1}H^TW^T \\
&= P^-_{k+1}-2P^-_{k+1}H^T\left(HP^-_{k+1}H^T\right)^{-1}HP^-_{k+1}+WHP^-_{k+1}H^TW^T \\
&= P^-_{k+1}-2P^-_{k+1}H^T\left(HP^-_{k+1}H^T\right)^{-1}\left(HP^-_{k+1}H^T\right)\left(HP^-_{k+1}H^T\right)^{-1}HP^-_{k+1}+WHP^-_{k+1}H^TW^T \\
&= P^-_{k+1}-2WHP^-_{k+1}H^TW^T+WHP^-_{k+1}H^TW^T
\end{aligned}
$$

Better State Observer Summary

System:

$$x(k+1) = Fx(k) + Gu(k) + v(k)$$
$$y(k) = Hx(k)$$

Observer

1. Predict

$$\hat{x}(k+1 \mid k) = F\hat{x}(k \mid k) + Gu(k)$$
$$P(k+1 \mid k) = FP(k \mid k)F^T + Q$$

2. Correction

$$W = P(k+1 \mid k)H\left(HP(k+1 \mid k)H^T \quad\right)^{-1}$$
$$\Delta x = W\left(y(k+1) - H\hat{x}(k+1 \mid k)\right)$$

3. Update

$$\hat{x}(k+1 \mid k+1) = \hat{x}(k+1 \mid k) + W\nu$$
$$P(k+1 \mid k+1) = P(k+1 \mid k) - WHP(k+1 \mid k)H^T W^T$$

•Note: there is a problem with the previous slide, namely the covariance matrix of the estimate P will be singular. This makes sense because with perfect sensor measurements the uncertainty in some directions will be zero. There is no uncertainty in the directions perpendicular to Ω