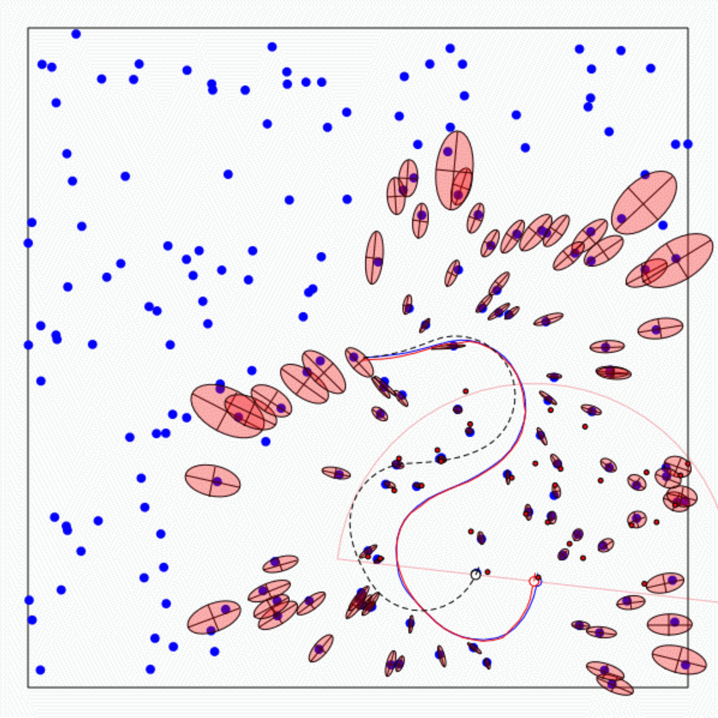
SLAM: Simultaneous Localization & Mapping

Given:

- Robot motion model: $\dot{x} = f(x, u) + \xi$
- The robot's controls, *u*
- Measurements (e.g., range, bearing) of nearby features: $y = h(x) + \omega$



Estimate:

- Map of landmarks (x_L)
- Robot's current *pose*, x_R, & its path
- Uncertainties in estimated quantities

$$\vec{x} = \begin{bmatrix}\vec{x}_R \\ \vec{x}_M\end{bmatrix} = \begin{bmatrix}\vec{x}_R \\ \vec{x}_{L_1} \\ \vdots \\ \vec{x}_{L_N}\end{bmatrix}$$

Robot state

Landmark positions

Estimation & Optimal (Kalman) Filtering

Observer

- Given $\overbrace{\dot{x} = f(x,u)}^{\text{Process Dynamics}}$ $\overbrace{y = h(x)}^{\text{Measurement Equation}}$
- Calculate, infer, deduce the state x from measurements y
- E.g. the *Luenberger Observer* $\dot{x} = Ax + Bu + L(y - Cx)$

Estimator

- Given $\dot{x} = f(x,u) + \xi$ $y = h(x) + \omega$
  - ξ represents *process noise/uncertainty* (e.g., gust or unmodeled effects)
  - ω represents *measurement noise/uncertainty*
- *Estimate* (in an *optimal)* way the state $x(t)$ based on
  - measurements $y(\tau)$, $\tau \leq t$
  - dynamic and measurement models
  - noise model(s).

The ***Kalman*** Filter

Applies rigorously to:

- Linear dynamical system

$$x_{k+1} = A_k x_k + B_k u_k + \xi_k$$

  - x_k= is the system state at time t_k (we want to estimate this)
  - u_k= control input at t_k (known)
  - $\xi_k$is the “process noise”
- Linear measurement equation:

$$y_{k+1}= H_k x_k + \omega_k$$

  - H_k= measurement matrix
  - ω_k= measurement noise
- Stochastic disturbances:
  - ξ and ω are *zero mean, white, Gaussian* random processes

Note: analogous theory for $x\dot(t) = Ax(t) + Bu(t) + \xi(t);\quad y(t) = Hx(t) + \omega(t)$

The *Kalman* Filter

The Kalman Filter (KF) aims to

- Find the “best” (min. variance) estimate of the state x at time t_k given
  - the dynamic and measurement models,
  - a characterization of the disturbances,
  - the measurements $y_0, y_1, \dots, y_k$
- Estimate the “uncertainty” in the state estimate.
- **Notation:** $\hat{x}_{k|j}$= the estimate of state x at t_k given measurements and information up until time t_j
  - $k = j \quad \Rightarrow$ KF state estimate
  - $k > j \quad \Rightarrow$ predictor
  - $k < j \quad \Rightarrow$ Kalman “smoother”

Recursive Structure of the KF

The KF has a ***2-step*** structure:

- Dynamic (time) update)

  “covariance” of the estimate

  - $\hat{x}_{k+1|k} = A_k \hat{x}_{k|k} + B_k u_k$
  - $\Sigma_{k+1|k} = A_k \Sigma_{k|k} A_k^T + B_k Q_k B_k^T$
- Measurement Update
  - $\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + \underbrace{K_{k+1}}_{\text{“Kalman Gain”}} \overbrace{\left(y_{k+1} - H_{k+1}\hat{x}_{k+1|k}\right)}^{\text{“residual,” “innovation”}}$
  - $\Sigma_{k+1|k+1} = \Sigma_{k+1|k} - \Sigma_{k+1|k} \mathrm{H}_{k+1}^{\mathrm{T}} \left(\mathrm{H}_{k+1} \Sigma_{k+1|k} \mathrm{H}_{k+1}^{\mathrm{T}} + \mathrm{R}_{k+1}\right)^{-1} \mathrm{H}_{k+1} \Sigma_{k+1|k}$

    $= (I - K_{k+1} H_{k+1}) \Sigma_{k+1|k} = \Sigma_{k+1|k} (\mathrm{I} - \mathrm{H}_{k+1}^{\mathrm{T}} \mathrm{K}_{k+1}^{\mathrm{T}})$

Where the “Kalman Gain” is:

- $K_{k+1} = \Sigma_{k+1|k} H_k^T \left(\underbrace{\mathrm{H}_{k+1} \Sigma_{k+1|k} \mathrm{H}_{k+1}^{\mathrm{T}}}_{\text{“How much do I trust the model?”}} + \underbrace{\mathrm{R}_{k+1}}_{\text{“How much do I trust the measurements?”}}\right)^{-1}$

Usefulness of the Kalman Filter

- Smooths (averages) noisy measurements, or multiple sensory inputs (sensor fusion)
- Recursive structure: next state estimate is only a function of the most recent measurement and previous estimate
  - Good for real time implementation
- Provides an *uncertainty measure* about its output.
  - Useful for fault detection
  - Make better decisions for autonomy

Let's derive the Kalman Filter in 3 steps:

- **Step #1:** assume ***NO*** process or measurement noise.
- **Step #2:** incorporate process noise
- **Step #3:** incorporate measurement noise

Case 1: No Noise

Dynamics & Measurement:

$$x_{k+1} = A_k x_k + B_k u_k; \qquad y_{k+1} = H_{k+1} x_{k+1}$$

- Assume that H_k is full rank
- Assume we are given $\hat{x}_{k|k}$ (recursive derivation)

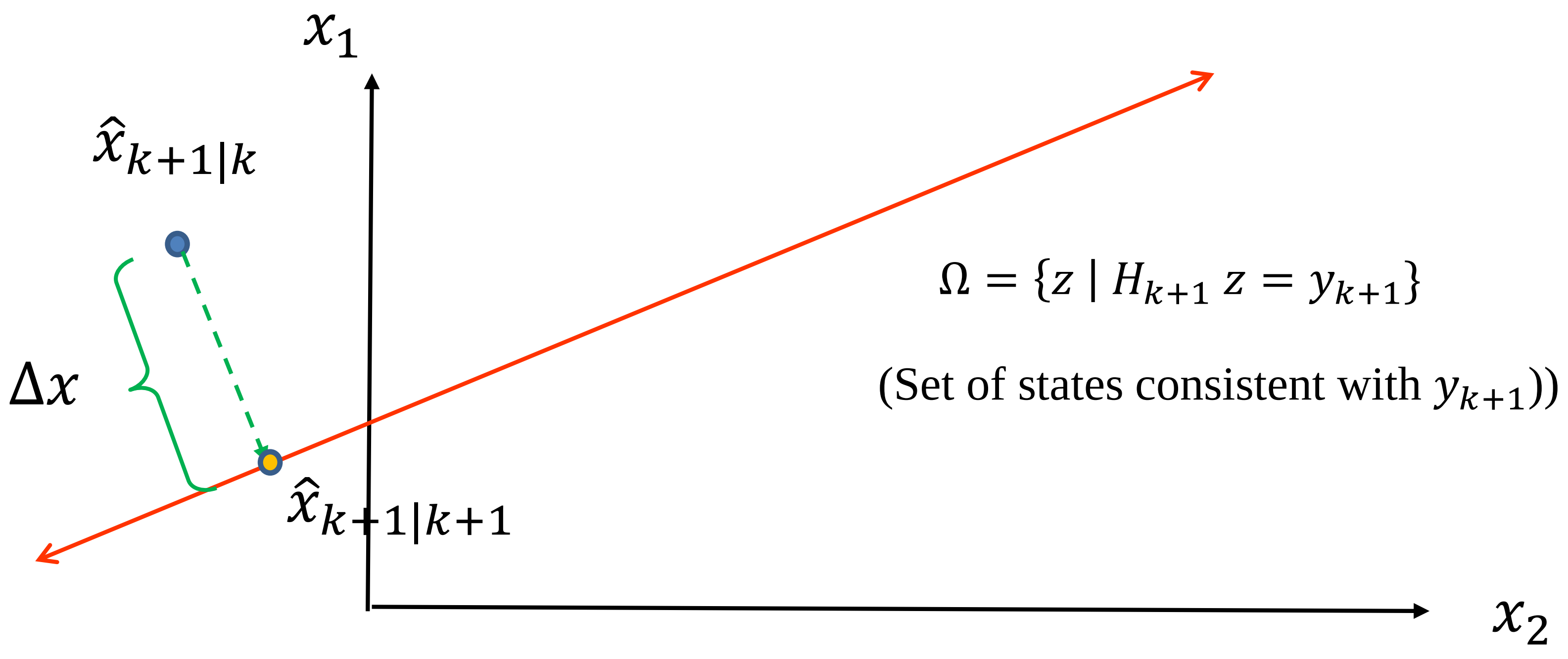
Prediction Step: $\hat{x}_{k+1|k} = A_k \, \hat{x}_{k|k} + B_k u_k$

Measurement Step: $y_{k+1} = H_{k+1} x_{k+1}$

- The true state at t_{k+1} must lie in Ω

$$\Omega = \{z \mid H_{k+1} z = y_{k+1}\}$$

- *Estimation principle*: the best state estimate is the point in Ω that is most consistent with the dynamical prediction.



$\Omega = \{z \mid H_{k+1}\, z = y_{k+1}\}$

(Set of states consistent with y_{k+1}))

- $\Delta x = \hat{x}_{k+1|k+1} - \hat{x}_{k+1|k}$
- For Δx to be the *shortest* vector (closest estimate in Ω, Δx must be $\perp$ to Ω
- **Solution #1:**
    - $\hat{x}_{k+1|k+1} = argmin\,(\Delta \mathrm{x})^2 \quad \mathrm{s.t.}\ \ \hat{x}_{k+1|k+1} \in \Omega$

Solution #1: Constrained Optimization

(not the textbook solution)

- Constrained Lagrangian: $L(x, \lambda) = (x - \hat{x})^2 + \lambda^T(y - Hx)$
  - Notation: $x \equiv \hat{x}_{k+1|k+1}; \quad \hat{x} \equiv \hat{x}_{k+1|k}; \quad y \equiv y_{k+1}$
  - Necessary Conditions: $\frac{\partial L}{\partial x} = \frac{\partial L}{\partial \lambda} = 0$
- $\frac{\partial L}{\partial x} = x - \hat{x} - H^T\lambda = 0 \;\; \Rightarrow \;\; x = \hat{x} + H^T\lambda \qquad (*)$
- $\frac{\partial L}{\partial \lambda} = y - Hx = 0 = y - H\left(\hat{x} + H^T\lambda\right) \;\; \Rightarrow$

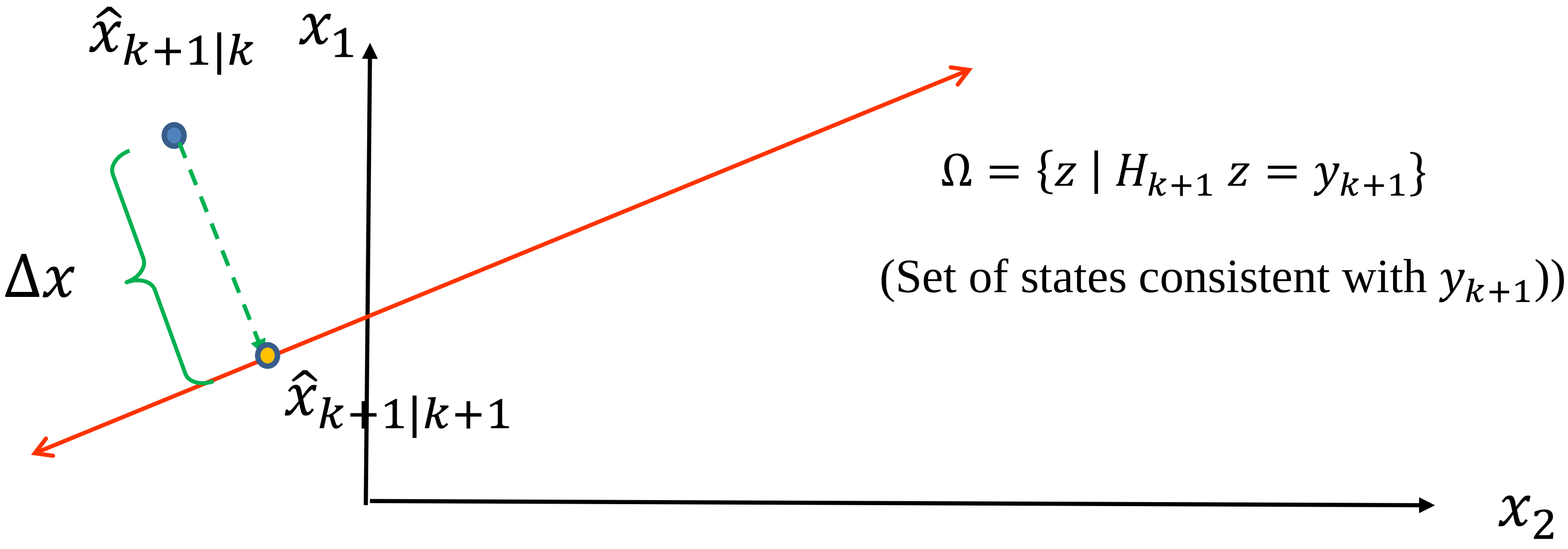
$$\lambda = \left(HH^T\right)^{-1}(y - H\hat{x})$$

- Substituting λ into (*) yields the result:

$$\hat{x}_{k+1|k+1} = \underbrace{\hat{x}_{k+1|k}}_{\text{Dynamic Prediction}} + {\underbrace{H_{k+1}^T\left(H_{k+1}H_{k+1}^T\right)}_{\text{Kalman “gain”}}}^{-1} \overbrace{\left(y_{k+1} - \underbrace{H_{k+1}\,\hat{x}_{k+1|k}}_{\text{Predicted measurement}}\right)}^{\text{innovation}}$$

Solution #2: Linear Algebra

(the textbook solution: parallels Kalman's method and later steps)



$\Omega = \{z \mid H_{k+1}\, z = y_{k+1}\}$

(Set of states consistent with y_{k+1}))

if $\Delta x \perp \Omega$, then $a^T \Delta x = 0$ for some a $\| \ \Omega$

- $a \ \| \ \Omega$ iff $H_{k+1} a = 0$ (i.e., a in $Null(H_{k+1})$)
- $b \perp Null(H_{k+1})$ if $b \in column(H_{k+1}^T)$
- Since $\Delta x \perp \Omega$, $\Delta x = H_{k+1}^T \gamma$ for some γ

Solution #2: Continued

- Innovation (again): $\nu = y_{k+1} - H_{k+1}\hat{x}_{k+1|k}$
- Assume $\gamma = K\,\nu$ (K is nearly the Kalman gain)
- Find K such that

$$y_{k+1} = H_{k+1}(\hat{x}_{k+1|k} + \Delta x)$$

  - That is, find $\Delta x = H_{k+1}^{T}\gamma = H_{k+1}^{T}K\nu$ that is $\perp\ \Omega$
- Rearrange (**):

$$H_{k+1}\Delta x = y_{k+1} - H_{k+1}\hat{x}_{k+1|k} = \nu$$

$$H_{k+1}\left(H_{k}^{T+1}K\nu\right) = \nu \quad \Rightarrow \quad K = \left(H_{k+1}H_{k+1}^{T}\right)^{-1}$$

- Putting all the pieces together:

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + \Delta x = \hat{x}_{k+1|k} + H_{k+1}\left(H_{k+1}H_{k+1}^{T}\right)^{-1}\nu$$

$$= \hat{x}_{k+1|k} + H_{k+1}\left(H_{k+1}H_{k+1}^{T}\right)^{-1}(y_{k+1} - H_{k+1}\,\hat{x}_{k+1|k})$$

Summary of Case 1: No Noise

Prediction Step: $\hat{x}_{k+1|k} = A_k\, \hat{x}_{k|k} + B_k u_k$

Measurement Step:

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + H_{k+1}^T \left(H_{k+1} H_{k+1}^T\right)^{-1} \left(y_{k+1} - H_{k+1}\, \hat{x}_{k+1|k}\right)$$

Practical Reality: This is ***not*** a good observer

- Not guaranteed to converge in all cases
- Assumes measurement relation is perfect
- Doesn’t correct well for errors parallel to Ω

Case 2: Noise in Dynamic Model

Dynamics & Measurement:

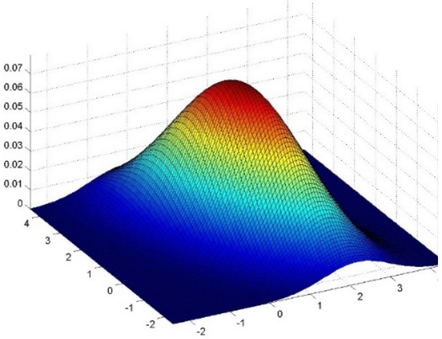
$$x_{k+1} = A_k x_k + B_k u_k + \xi; \qquad y_{k+1} = H_{k+1} x_{k+1}$$

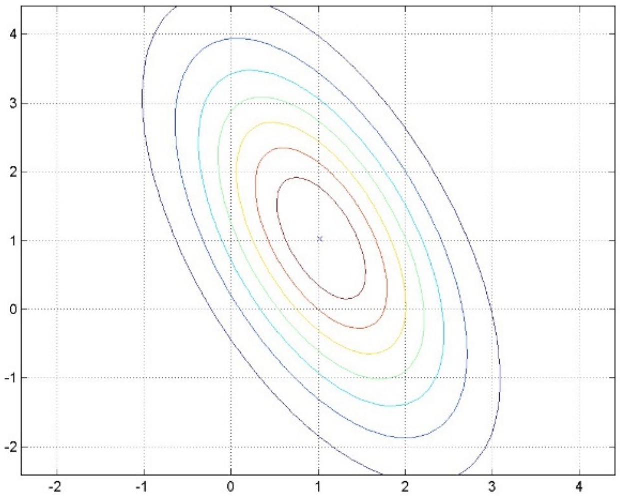
- H_k is full rank and $\hat{x}_{k|k}$ given

Prediction Step: $\hat{x}_{k+1|k} = A_k\, \hat{x}_{k|k} + B_k u_k$

- But $\hat{x}_{k+1|k}$ is now a ***random variable*** *(zero mean Gaussian)*

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





Case 2: Noise in Dynamic Model

Covariance of Dynamic prediction:

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

Substitute $x_{k+1} = A_k\, x_{k|k} + B_k u_k + \xi_k$

$$\hat{x}_{k+1|k} = A_k\, \hat{x}_{k|k} + B_k u_k$$

- Use fact that expectation is linear and that $x_k$and $\hat{x}_{k|k}$ are independent of ξ_k
- $P_{k+1|k} = A_k P_{k|k} A_k^T + V_k$

Measurement Step: $y_{k+1} = H_{k+1} x_{k+1}$

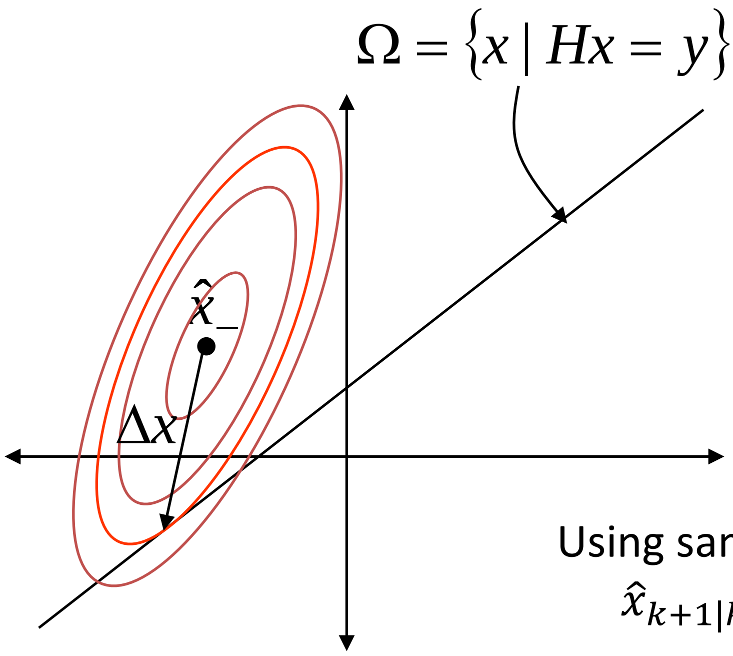
- The true state at t_{k+1} must lie in Ω

$$\Omega = \{z \mid H_{k+1} z = y_{k+1}\}$$

Finding the correction (geometric intuition)

Given the prediction $\hat{x}_{k+1|k}$, the covariance $P_{k+1|k}$ and the measurement y_{k+1}, find Δx so that $\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + \Delta \mathrm{x}$ is the *best* or *most probable* estimate. I.

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



The most probable Δx is the one that

- minimizes $\Delta x^T P_{k+1|k}^{-1} \Delta x$ (since this maximizes $p(x)$)
- while $x_{k+1|k+1} \in \Omega$
- $\Delta x^T P_{k+1|k}^{-1} \Delta x$ is the Mahalonobis Distance

Using same analysis as Solution 1 or 2 of Case 1, yields:

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}[y_{k+1} - H_{k+1}\,\hat{x}_{k+1|k}]$$

$$K_{k+1} = P_{k+1|k} H_{k+1}^T \left[H_{k+1}\,P_{k+1|k}\,H_{k+1}^T\right]^{-1}$$

Case 3: Add Measurement Noise

Dynamics & Measurement:

$$x_{k+1} = A_k x_k + B_k u_k + \xi; \qquad y_{k+1} = H_{k+1} x_{k+1} + \omega_{k+1}$$

- H_k is full rank, $\omega_{k+1} \sim N(0, R_{k+1})$ and $\hat{x}_{k|k}$ given

Prediction Step: Same as Step #2

- $\hat{x}_{k+1|k} = A_k \, \hat{x}_{k|k} \; + B_k u_k$
- $P_{k+1|k} = A_k P_{k|k} A_k^T + V_k$

Measurement Update:

- No longer a "crisp" observation constraint. We only know that the output is draw from a Gaussian distribution centered at y_{k+1}
- Look for ***most likely output***

Finding the correction (with output noise)

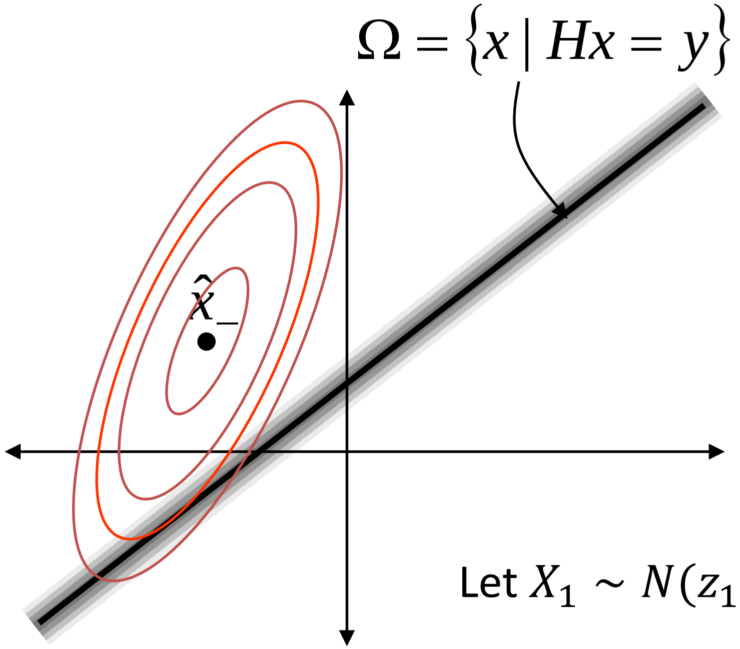
$$y = Hx + w$$

Since you don't have a hyperplane to aim for, you can't solve this with algebra!

You have to solve an optimization problem

The most likely output is the *most likely* point

- the measurement $y_{k+1} = H_{k+1}x_{k+1} + \omega_{k+1}$ is the sum of two ***independent*** random variables.
- most likely variable maximizes the joint probability of these two random variables, which is their product.



Let $X_1 \sim N(z_1, C_1)$ and $X_1 \sim N(z_2, C_2)$. Then $p(x_1)p(x_2) =$

$$N(z_1 + C_1(C_1 + C_2)^{-1}(z_2 - z_1), C_1 - C_1(C_1 + C_2)^{-1}C_1$$

Recursive Structure of the KF

The KF has a ***2-step*** structure:

- Dynamic (time) update)
  - $\hat{x}_{k+1|k} = A_k\hat{x}_{k|k} + B_k u_k$
  - $\Sigma_{k+1|k} = A_k\Sigma_{k|k}A_k^T + V_k$

“covariance” of the estimate

- Measurement Update
  - $\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + \underbrace{K_{k+1}}_{\text{“Kalman Gain”}}\overbrace{(y_{k+1} - H_{k+1}\hat{x}_{k+1|k})}^{\text{“residual,” “innovation”}}$
  - $\Sigma_{k+1|k+1} = \Sigma_{k+1|k} - \Sigma_{k+1|k}\mathrm{H}_{k+1}^{\mathrm{T}}\left(\mathrm{H}_{k+1}\Sigma_{k+1|k}\mathrm{H}_{k+1}^{\mathrm{T}} + \mathrm{R}_{k+1}\right)^{-1}\mathrm{H}_{k+1}\Sigma_{k+1|k}$

$$= (I - K_{k+1}H_{k+1})\Sigma_{k+1|k} = \Sigma_{k+1|k}(\mathrm{I} - \mathrm{H}_{k+1}^{\mathrm{T}}\mathrm{K}_{k+1}^{\mathrm{T}})$$

Where the “Kalman Gain” is:

- $K_{k+1} = \Sigma_{k+1|k}H_k^T\left(\underbrace{\mathrm{H}_{k+1}\Sigma_{k+1|k}\mathrm{H}_{k+1}^{\mathrm{T}}}_{\text{“How much do I trust the model?”}} + \underbrace{\mathrm{R}_{k+1}}_{\text{“How much do I trust the measurements?”}}\right)^{-1}$