The Extended Kalman Filter

D. A. Forsyth, UIUC

The Kalman filter is wonderful, but…

- The linear model of measurement isn't always helpful
  - Example:
    - our localization procedures
- The linear model of movement isn't always helpful
  - Example:
    - the car in ND Kalman filter example is completely unrealistic

- What to do about non-linearities?

Example: Hard localization

- Example we saw
  - Assume
    - car state is (position; velocity; acceleration)
      - it doesn’t rotate!
      - this yields D_i, and noise
    - we know M_i and noise for beacons
  - Then it’s all easy (plug in equations and go)

- but what if we localize with IRLS?

IRLS localization

- Write state of vehicle: $$\mathbf{x}_i = \left(\begin{array}{c} \text{position} \\ \text{velocity} \end{array} \right)$$
- Can extract position as: $$\mathbf{p}_i = \Pi_p \mathbf{x}_i$$
- State update is: $$\mathbf{x}_i = \mathcal{D}_i \mathbf{x}_{i-1} + \xi$$
- Measurement is:

$$\mathbf{y}_i = \text{argmin}_{\mathbf{u}} C(\mathbf{u}, \mathbf{x}_i)$$

HUH?

Harder localization trick

- Model the cost function as:

$$C(\mathbf{u}, \mathbf{x}_i) \approx c_0 + \mathbf{v}^T(\mathbf{u} - \mathbf{p}_i) + (\mathbf{u} - \mathbf{p}_i)^T \frac{\mathcal{H}}{2}(\mathbf{u} - \mathbf{p}_i)$$

  - at the minimum - so actually, v=0
  - now the cost function might be slightly wrong, which will cause errors in u
  - if we use the model:

$$\mathbf{y}_i = \mathbf{u} = \mathbf{p}_i + \mathcal{H}^{-1/2}\zeta$$

  - then we have:

$$C(\mathbf{y}_i, \mathbf{x}_i) \sim N(c_0, \frac{d\sigma^2}{2})$$

$$N(0, \sigma^2 I)$$

And a kind of "evenness" property

Harder localization trick, II

- This property is reasonable:
    - we can’t tell noise directions apart by their effect on the cost function

- Now we’re in business:

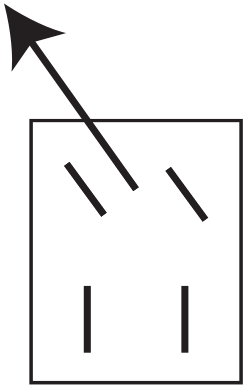
$$\mathbf{x}_i \sim N(\mathcal{D}_i \mathbf{x}_{i-1}, \Sigma_i)$$

$$\mathbf{y}_i \sim N(\Pi_p \mathbf{x}_i, \sigma^2 \mathcal{H}_i^{-1})$$

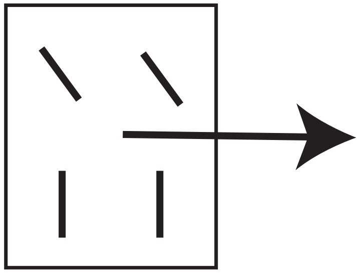
Choose this

Hessian of cost function at best location

Example: Nasty dynamical model



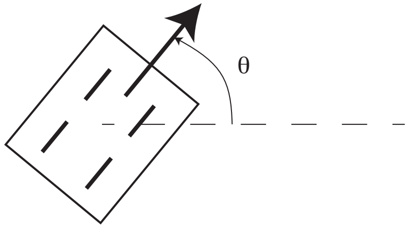
OK



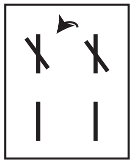
Not OK

Formally: car is non-holonomic

Building a movement model

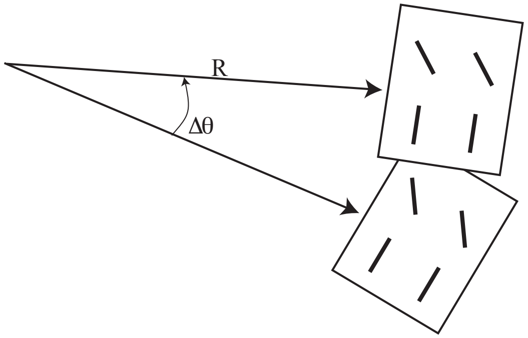


$$\begin{bmatrix} x \\ y \\ \theta \end{bmatrix} \rightarrow \begin{bmatrix} x + v\Delta t cos\theta \\ y + v\Delta t \sin\theta \\ \theta \end{bmatrix}$$



$$\begin{bmatrix} x \\ y \\ \theta \end{bmatrix} \rightarrow \begin{bmatrix} x \\ y \\ \theta + \Delta\theta \end{bmatrix}$$

A general movement model



$$\begin{bmatrix} x \\ y \\ \theta \end{bmatrix} \rightarrow \begin{bmatrix} x + R(\sin(\theta + \Delta\theta) - \sin\theta) \\ y - R(\cos(\theta + \Delta\theta) - \cos\theta) \\ \theta + \Delta\theta \end{bmatrix}$$

THIS ISN'T LINEAR!

For sufficiently small timestep, bounded rate of change in angle, we get

$$\begin{bmatrix} x \\ y \\ \theta \end{bmatrix} \rightarrow \begin{bmatrix} x + v\cos\theta \\ y + v\sin\theta \\ \theta + u \end{bmatrix}$$

v, u parameters of motion

THIS ISN'T LINEAR!

The extended Kalman filter

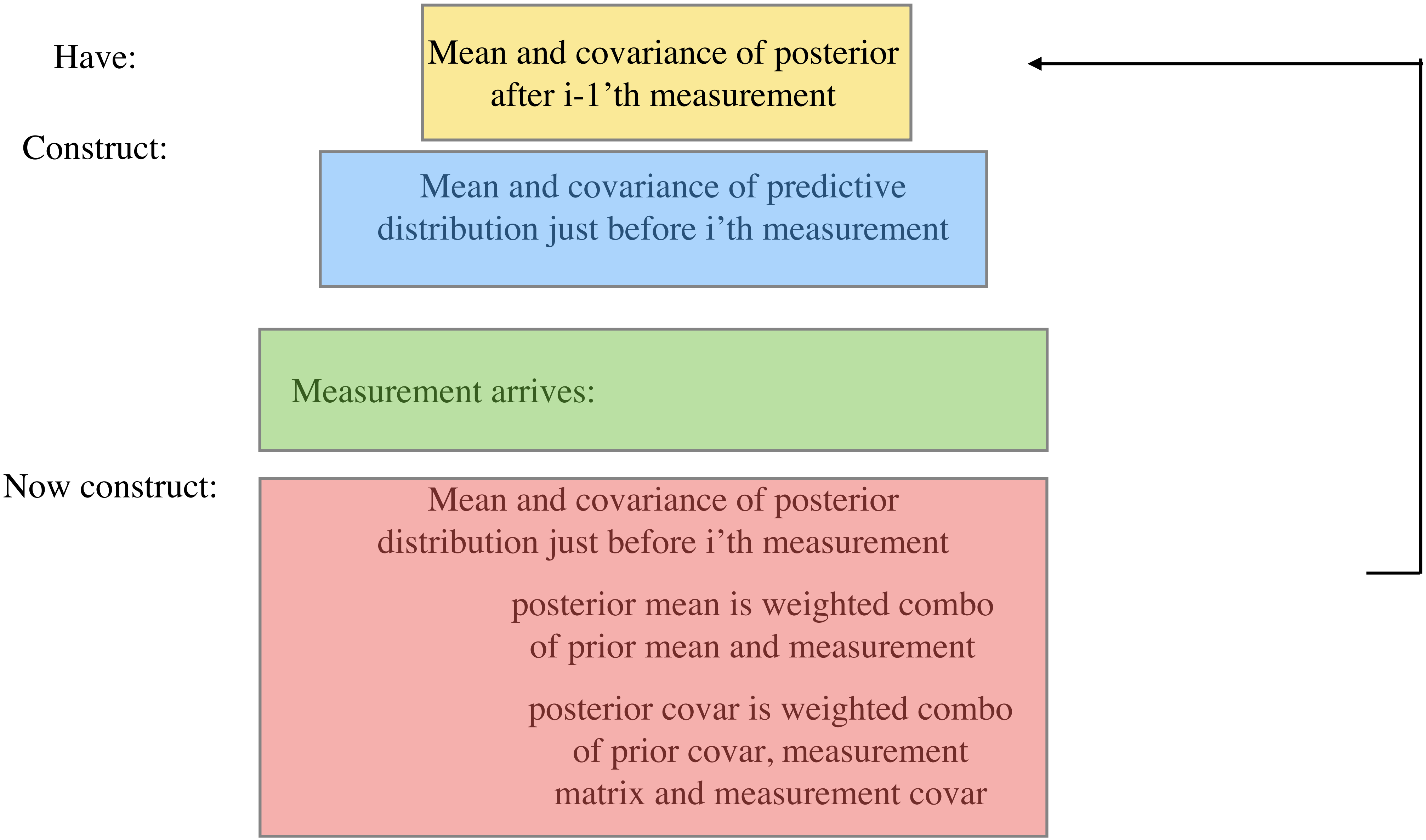
- What happens if state update, measurement aren’t linear?
  - particle filter
  - linearize and approximate (EKF)

$$\mathbf{x}_i = f(\mathbf{x}_{i-1}, \underset{\uparrow}{\mathbf{n}})$$

Noise - normal, mean 0, Cov known

$$\mathbf{y}_i = g(\mathbf{x}_i, \overset{\downarrow}{\mathbf{n}})$$

The steps, KF:



Have:

Mean and covariance of posterior after i-1’th measurement

Construct:

Mean and covariance of predictive distribution just before i’th measurement

Measurement arrives:

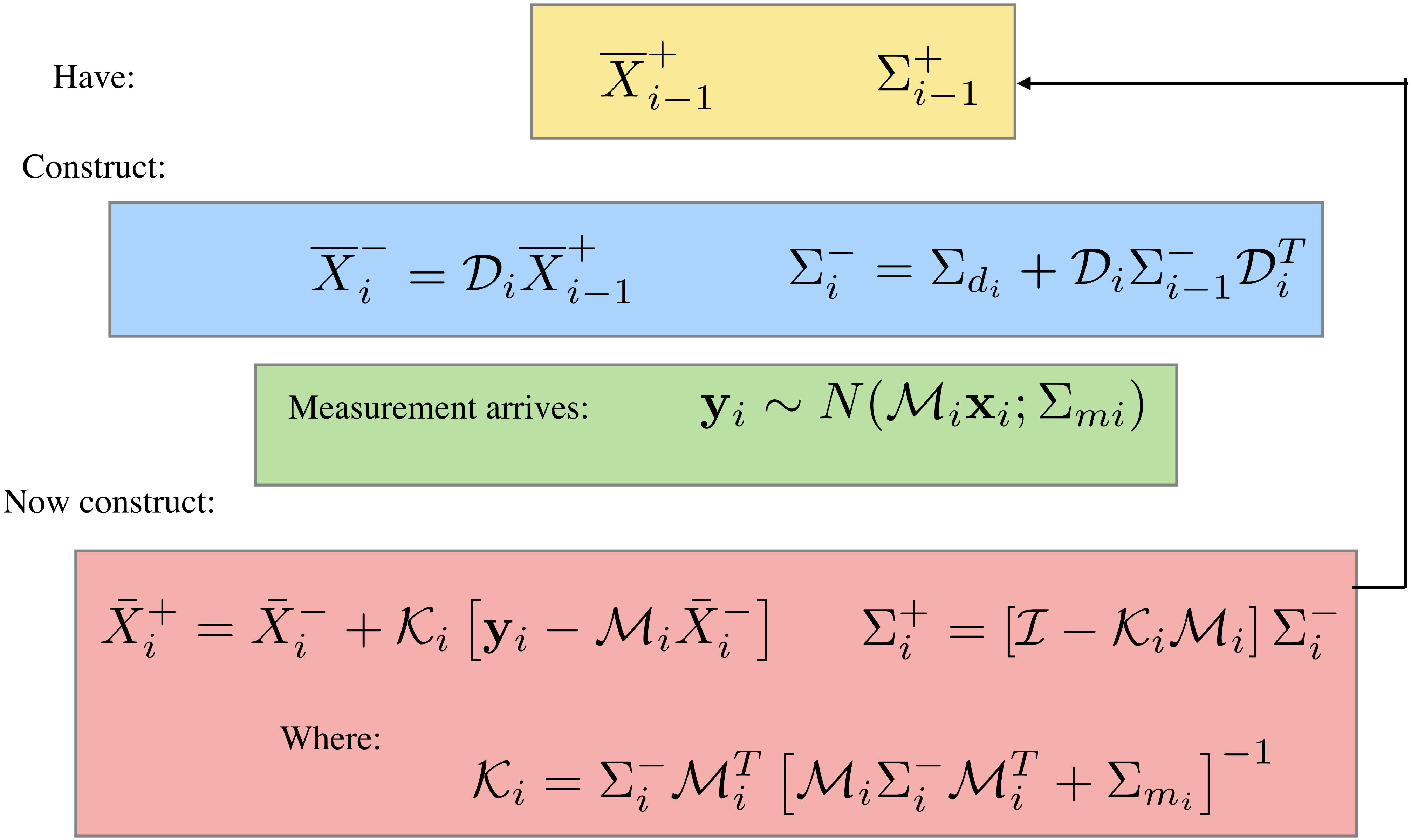
Now construct:

Mean and covariance of posterior distribution just before i’th measurement

posterior mean is weighted combo of prior mean and measurement

posterior covar is weighted combo of prior covar, measurement matrix and measurement covar

The steps, KF:



Have:

$$\overline{X}^{+}_{i-1} \qquad \Sigma^{+}_{i-1}$$

Construct:

$$\overline{X}^{-}_{i} = \mathcal{D}_i \overline{X}^{+}_{i-1} \qquad \Sigma^{-}_{i} = \Sigma_{d_i} + \mathcal{D}_i \Sigma^{-}_{i-1} \mathcal{D}_i^T$$

Measurement arrives: $\mathbf{y}_i \sim N(\mathcal{M}_i \mathbf{x}_i; \Sigma_{mi})$

Now construct:

$$\bar{X}^{+}_{i} = \bar{X}^{-}_{i} + \mathcal{K}_i \left[\mathbf{y}_i - \mathcal{M}_i \bar{X}^{-}_{i}\right] \qquad \Sigma^{+}_{i} = \left[\mathcal{I} - \mathcal{K}_i \mathcal{M}_i\right] \Sigma^{-}_{i}$$

Where:

$$\mathcal{K}_i = \Sigma^{-}_{i} \mathcal{M}_i^T \left[\mathcal{M}_i \Sigma^{-}_{i} \mathcal{M}_i^T + \Sigma_{m_i}\right]^{-1}$$

Linearization and noise

- Two ways in which noise could affect x_i
    - x_{i-1} is noisy
    - AND there is n to account for

$$\mathbf{x}_i = f(\mathbf{x}_{i-1}, \overset{\downarrow}{\mathbf{n}})$$

- Now consider some nonlinear function with noisy input
    - first case

$$h(\mathbf{x}) \text{ where } \mathbf{x} \sim N(\bar{\mathbf{x}}, \Sigma_x)$$

$$h(\bar{\mathbf{x}} + \zeta) \text{ where } \zeta \sim N(0, \Sigma_x)$$

Approximate

$$h(\bar{\mathbf{x}} + \zeta) \approx h(\bar{\mathbf{x}}) + J_{h,x}\zeta$$

Yields

$$h(\mathbf{x}) \sim N(h(\bar{\mathbf{x}}), J_{h,x}\Sigma_x J_{h,x}^T)$$

$$J_{h,x} = \left[\begin{array}{ccc} \frac{\partial h_1}{\partial x_1} & \cdots & \cdots \\ \cdots & \frac{\partial h_i}{\partial x_j} & \cdots \\ \cdots & \cdots & \cdots \end{array}\right]$$

Jacobian === derivative

Linearization and noise

- Two ways in which noise could affect x_i
  - x_{i-1} is noisy
  - AND there is n to account for

$$\mathbf{x}_i = f(\mathbf{x}_{i-1}, \overset{\downarrow}{\mathbf{n}})$$

- Now consider some nonlinear function with fixed input, noise
  - second case

$$h(\mathbf{x}, \mathbf{n}) \text{ where } \mathbf{n} \sim N(0, \sigma_n)$$

Approximate

$$h(\mathbf{x}, \mathbf{n}) \approx h(\mathbf{x}, \mathbf{0}) + J_{h,n}\mathbf{n}$$

$$J_{h,n} = \begin{bmatrix} \frac{\partial h_1}{\partial n_1} & \cdots & \cdots \\ \cdots & \frac{\partial h_i}{\partial n_j} & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}$$

Yields

$$h(\mathbf{x}, \mathbf{n}) \sim N(h(\mathbf{x}, \mathbf{0}), J_{h,n}\Sigma_n J_{h,n}^T)$$

Jacobian === derivative

The extended Kalman filter

- Linearize:

$$\mathbf{x}_i = f_i(\mathbf{x}_{i-1}, \mathbf{n})$$

$$\mathcal{F}_x = \left[\begin{array}{ccc} \frac{\partial f_1}{\partial x_1} & \cdots & \cdots \\ \cdots & \frac{\partial f_i}{\partial x_j} & \cdots \end{array} \right]$$

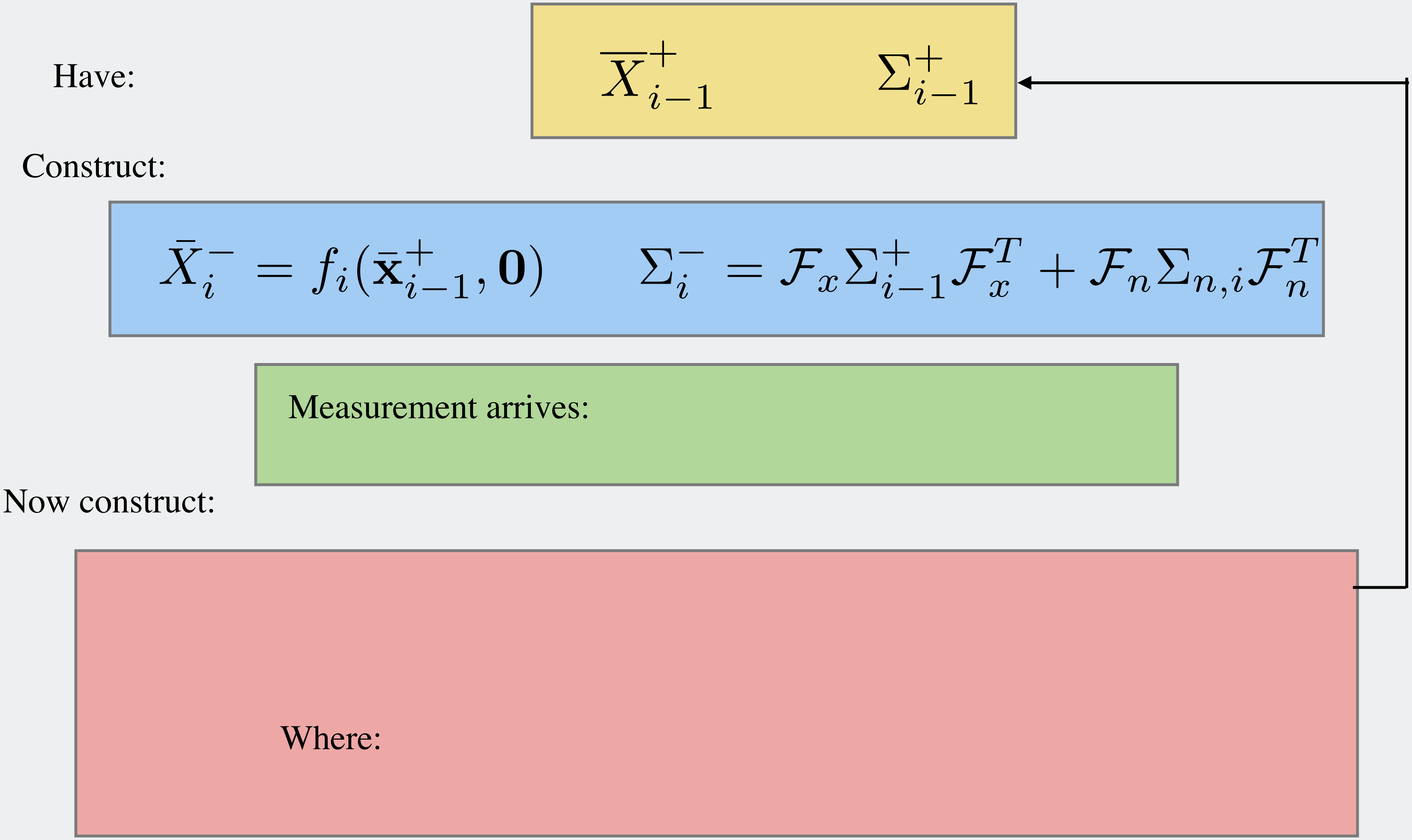
$$\mathcal{F}_n = \left[\begin{array}{ccc} \frac{\partial f_1}{\partial n_1} & \cdots & \cdots \\ \cdots & \frac{\partial f_i}{\partial n_j} & \cdots \end{array} \right]$$

Posterior covariance of x_{i-1}

$$\mathbf{x}_i \sim N(f_i(\bar{\mathbf{x}}^+_{i-1}, \mathbf{0}), \mathcal{F}_x \Sigma^+_{i-1} \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T)$$

Noise covariance

The steps, EKF:



Have:

$$\overline{X}_{i-1}^{+} \qquad \Sigma_{i-1}^{+}$$

Construct:

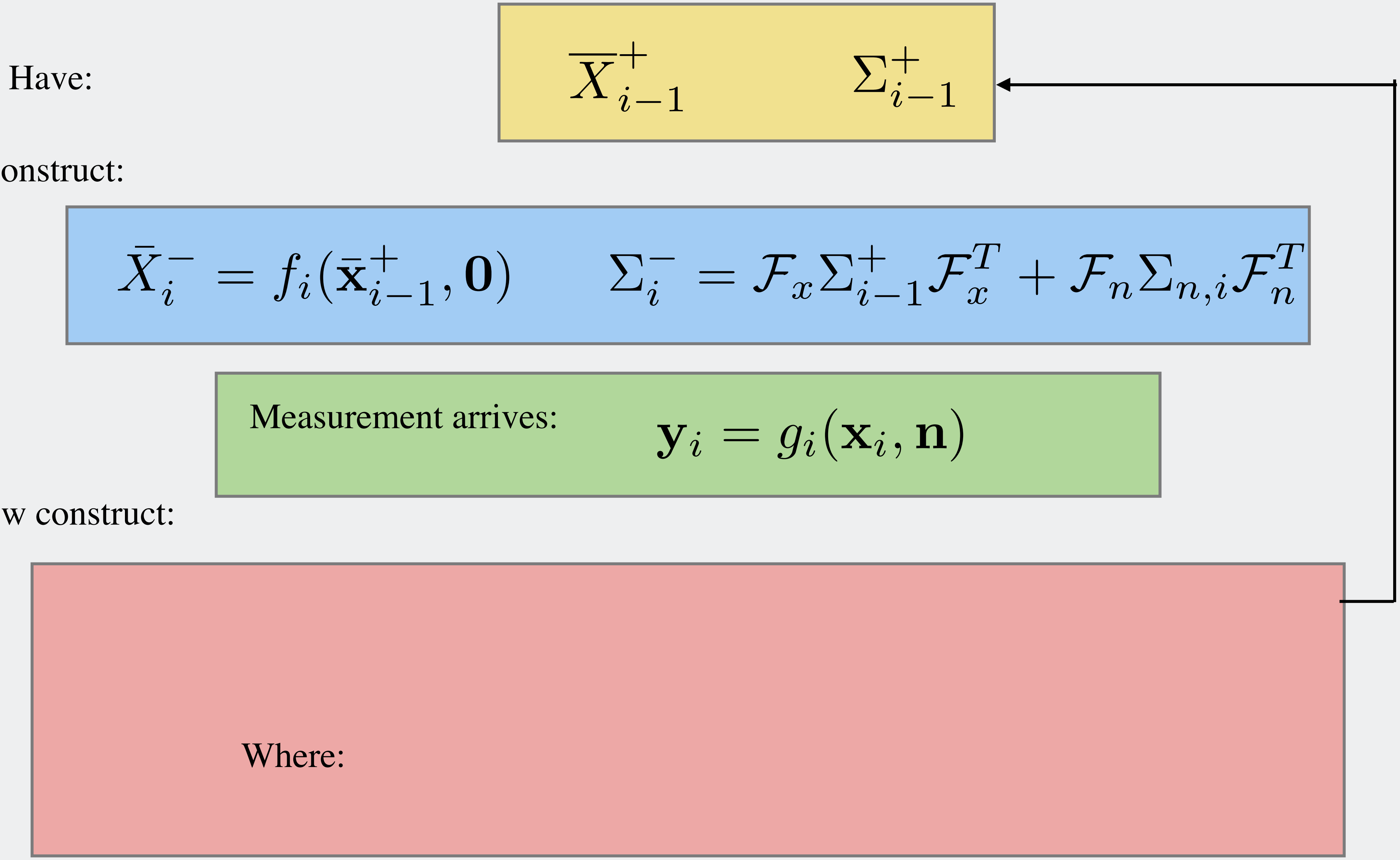
$$\bar{X}_i^- = f_i(\bar{\mathbf{x}}_{i-1}^+, \mathbf{0}) \qquad \Sigma_i^- = \mathcal{F}_x \Sigma_{i-1}^+ \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

Measurement arrives:

Now construct:

Where:

The steps, EKF:



Have:

$$\overline{X}^+_{i-1} \qquad \Sigma^+_{i-1}$$

Construct:

$$\bar{X}_i^- = f_i(\bar{\mathbf{x}}^+_{i-1}, \mathbf{0}) \qquad \Sigma_i^- = \mathcal{F}_x \Sigma^+_{i-1} \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

Measurement arrives: $\mathbf{y}_i = g_i(\mathbf{x}_i, \mathbf{n})$

Now construct:

Where:

The extended Kalman filter

- Linearize:

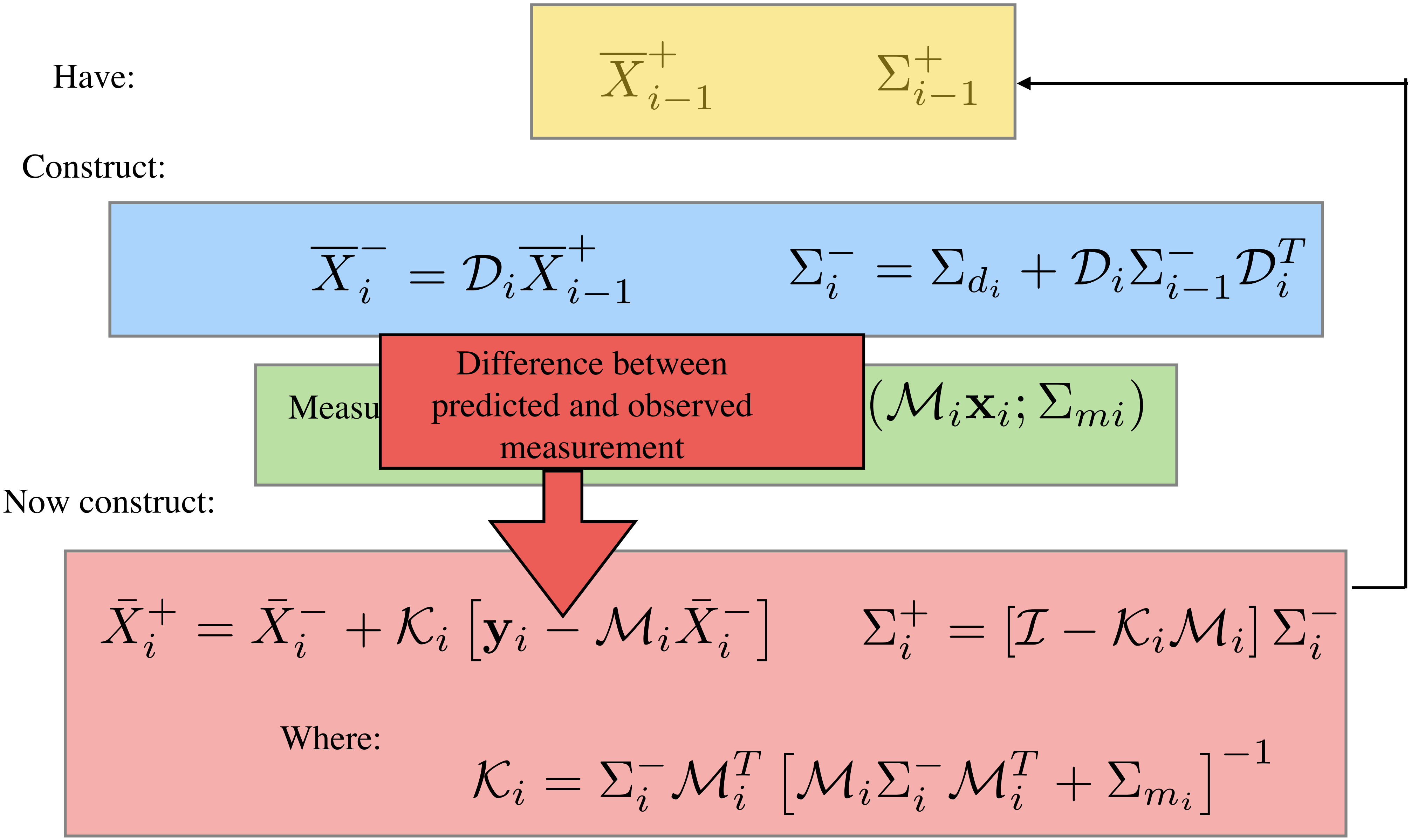
$$\mathbf{y}_i = g_i(\mathbf{x}_i, \mathbf{n})$$

$$\mathcal{G}_x = \left[\begin{array}{ccc} \frac{\partial g}{\partial x_1} & \cdots & \cdots \\ \cdots & \frac{\partial g}{\partial x_1} & \cdots \end{array} \right]$$

$$\mathcal{G}_n = \left[\begin{array}{ccc} \frac{\partial g}{\partial n_1} & \cdots & \cdots \\ \cdots & \frac{\partial g}{\partial n_1} & \cdots \end{array} \right]$$

$$\mathbf{y}_i \approx \mathcal{N}(g_i(\bar{X}_i^-, \mathbf{0}), \mathcal{G}_x \Sigma_i^- \mathcal{G}_x^T + \mathcal{G}_n \Sigma_{m,i} \mathcal{G}_n^T)$$

Recall: The steps, KF:



Have:

$$\overline{X}^+_{i-1} \qquad \Sigma^+_{i-1}$$

Construct:

$$\overline{X}^-_i = \mathcal{D}_i \overline{X}^+_{i-1} \qquad \Sigma^-_i = \Sigma_{d_i} + \mathcal{D}_i \Sigma^-_{i-1} \mathcal{D}_i^T$$

Measu $(\mathcal{M}_i \mathbf{x}_i; \Sigma_{mi})$

Difference between predicted and observed measurement

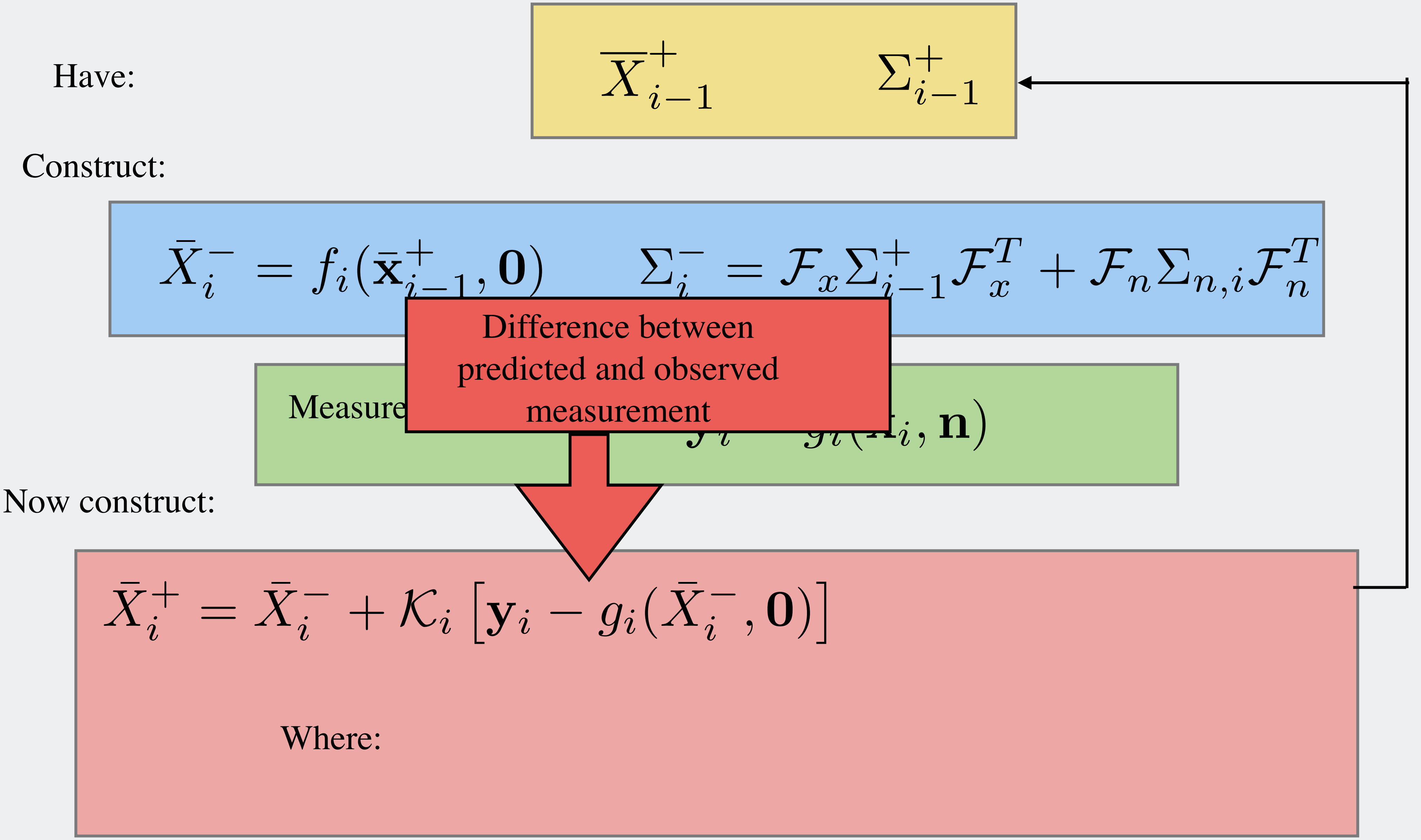
Now construct:

$$\bar{X}^+_i = \bar{X}^-_i + \mathcal{K}_i \left[\mathbf{y}_i - \mathcal{M}_i \bar{X}^-_i\right] \qquad \Sigma^+_i = \left[\mathcal{I} - \mathcal{K}_i \mathcal{M}_i\right] \Sigma^-_i$$

Where:

$$\mathcal{K}_i = \Sigma^-_i \mathcal{M}_i^T \left[\mathcal{M}_i \Sigma^-_i \mathcal{M}_i^T + \Sigma_{m_i}\right]^{-1}$$

The steps, EKF:



Have:

$$\overline{X}_{i-1}^{+} \qquad \Sigma_{i-1}^{+}$$

Construct:

$$\bar{X}_i^- = f_i(\bar{\mathbf{x}}_{i-1}^+, \mathbf{0}) \qquad \Sigma_i^- = \mathcal{F}_x \Sigma_{i-1}^+ \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

Difference between predicted and observed measurement

Measure

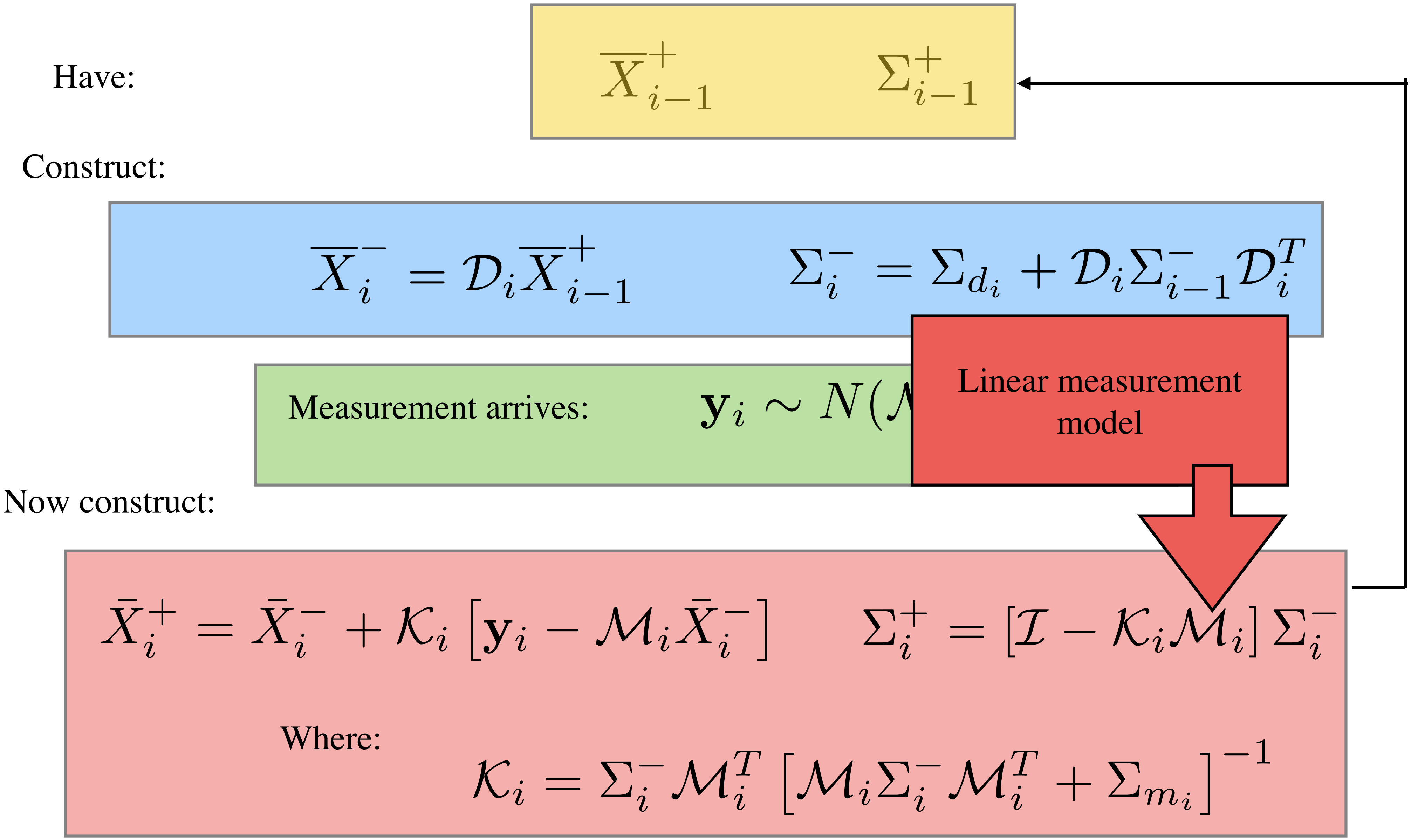
$$\mathbf{y}_i \quad g_i(\mathbf{x}_i, \mathbf{n})$$

Now construct:

$$\bar{X}_i^+ = \bar{X}_i^- + \mathcal{K}_i \left[\mathbf{y}_i - g_i(\bar{X}_i^-, \mathbf{0}) \right]$$

Where:

Recall: The steps, KF:



Have:

$$\overline{X}^{+}_{i-1} \qquad \Sigma^{+}_{i-1}$$

Construct:

$$\overline{X}^{-}_{i} = \mathcal{D}_i \overline{X}^{+}_{i-1} \qquad \Sigma^{-}_{i} = \Sigma_{d_i} + \mathcal{D}_i \Sigma^{-}_{i-1} \mathcal{D}_i^{T}$$

Measurement arrives: $\mathbf{y}_i \sim N($

Linear measurement model

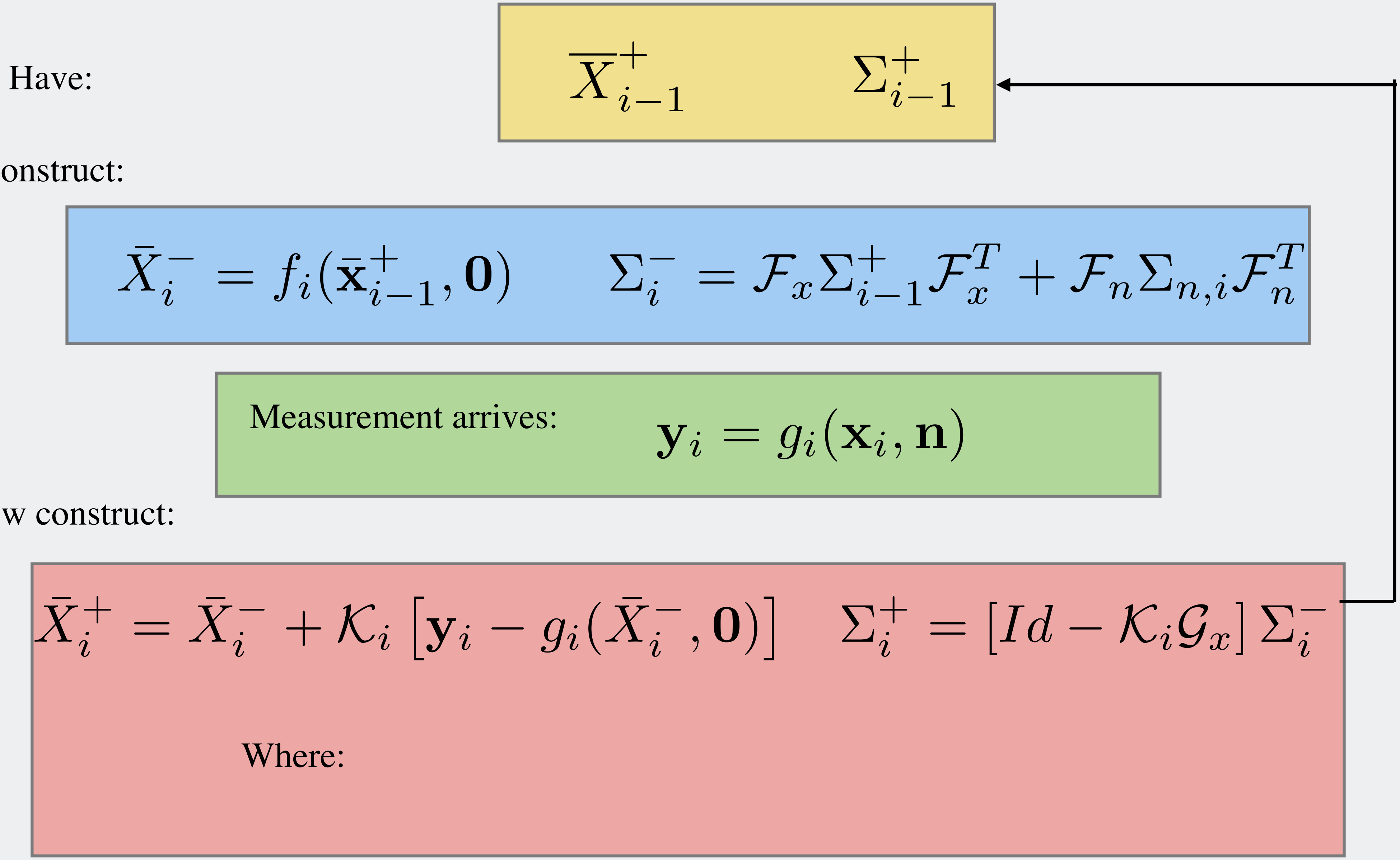
Now construct:

$$\bar{X}_i^{+} = \bar{X}_i^{-} + \mathcal{K}_i \left[\mathbf{y}_i - \mathcal{M}_i \bar{X}_i^{-} \right] \qquad \Sigma_i^{+} = \left[\mathcal{I} - \mathcal{K}_i \mathcal{M}_i \right] \Sigma_i^{-}$$

Where:

$$\mathcal{K}_i = \Sigma_i^{-} \mathcal{M}_i^{T} \left[\mathcal{M}_i \Sigma_i^{-} \mathcal{M}_i^{T} + \Sigma_{m_i} \right]^{-1}$$

The steps, EKF:



Have:

$$\overline{X}^{+}_{i-1} \qquad \Sigma^{+}_{i-1}$$

Construct:

$$\bar{X}_i^{-} = f_i(\bar{\mathbf{x}}^{+}_{i-1}, \mathbf{0}) \qquad \Sigma_i^{-} = \mathcal{F}_x \Sigma^{+}_{i-1} \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

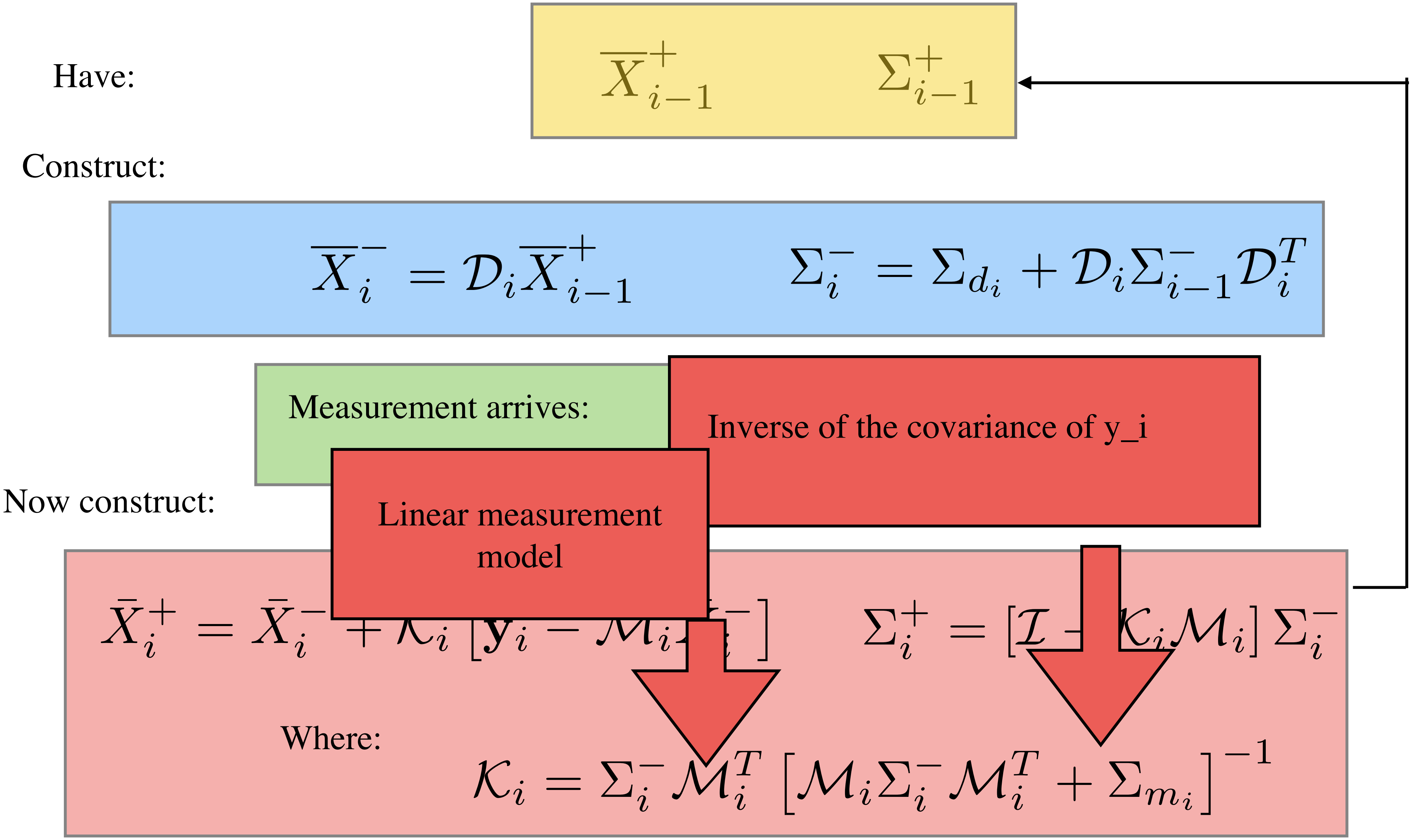
Measurement arrives: $\mathbf{y}_i = g_i(\mathbf{x}_i, \mathbf{n})$

Now construct:

$$\bar{X}_i^{+} = \bar{X}_i^{-} + \mathcal{K}_i \left[\mathbf{y}_i - g_i(\bar{X}_i^{-}, \mathbf{0})\right] \qquad \Sigma_i^{+} = \left[Id - \mathcal{K}_i \mathcal{G}_x\right] \Sigma_i^{-}$$

Where:

Recall: The steps, KF:



Have:

$$\overline{X}_{i-1}^{+} \qquad \Sigma_{i-1}^{+}$$

Construct:

$$\overline{X}_i^{-} = \mathcal{D}_i \overline{X}_{i-1}^{+} \qquad \Sigma_i^{-} = \Sigma_{d_i} + \mathcal{D}_i \Sigma_{i-1}^{-} \mathcal{D}_i^T$$

Measurement arrives:

Inverse of the covariance of y_i

Now construct:

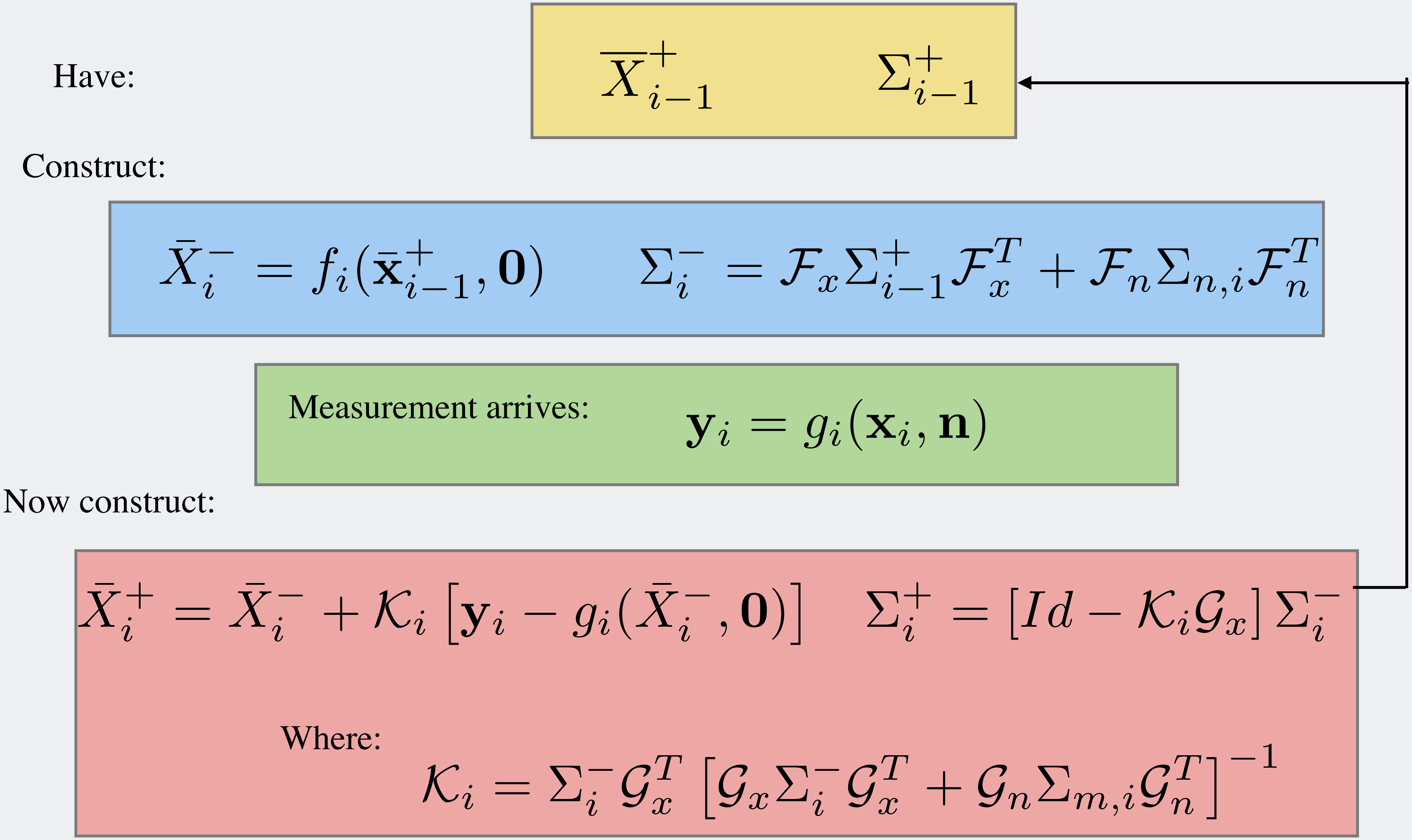
Linear measurement model

$$\bar{X}_i^{+} = \bar{X}_i^{-} + \mathcal{K}_i \left[\mathbf{y}_i - \mathcal{M}_i \bar{X}_i^{-}\right] \qquad \Sigma_i^{+} = \left[\mathcal{I} - \mathcal{K}_i \mathcal{M}_i\right] \Sigma_i^{-}$$

Where:

$$\mathcal{K}_i = \Sigma_i^{-} \mathcal{M}_i^T \left[\mathcal{M}_i \Sigma_i^{-} \mathcal{M}_i^T + \Sigma_{m_i}\right]^{-1}$$

The steps, EKF:



Have:

$$\overline{X}^{+}_{i-1} \qquad \Sigma^{+}_{i-1}$$

Construct:

$$\bar{X}_i^{-} = f_i(\bar{\mathbf{x}}^{+}_{i-1}, \mathbf{0}) \qquad \Sigma_i^{-} = \mathcal{F}_x \Sigma^{+}_{i-1} \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

Measurement arrives: $\mathbf{y}_i = g_i(\mathbf{x}_i, \mathbf{n})$

Now construct:

$$\bar{X}_i^{+} = \bar{X}_i^{-} + \mathcal{K}_i \left[\mathbf{y}_i - g_i(\bar{X}_i^{-}, \mathbf{0})\right] \qquad \Sigma_i^{+} = \left[Id - \mathcal{K}_i \mathcal{G}_x\right] \Sigma_i^{-}$$

Where:

$$\mathcal{K}_i = \Sigma_i^{-} \mathcal{G}_x^T \left[\mathcal{G}_x \Sigma_i^{-} \mathcal{G}_x^T + \mathcal{G}_n \Sigma_{m,i} \mathcal{G}_n^T\right]^{-1}$$

Outcome and issues

- In principle, can now filter position/orientation wrt map
  - linearize dynamics following recipe above
  - linearize measurements ditto
- There could be problems
  - EKF’s are fine if the linearization is reliable
    - can be awful if not (examples)
  - in fact, the map points are uncertain
    - why not try to make/update map while moving? SLAM, to follow