Filtro de Kalman Estendido

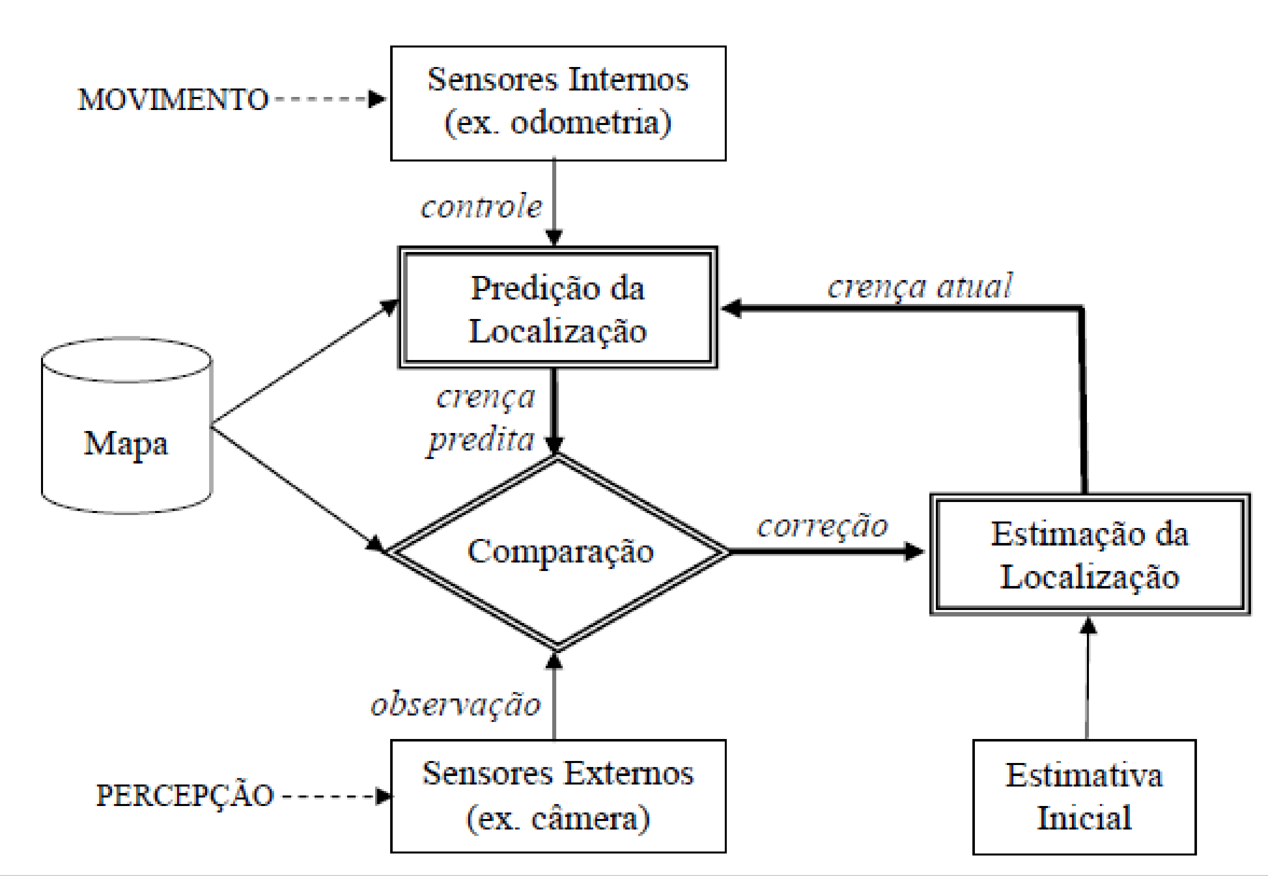
Valéria de Carvalho Santos
Prof. Roseli A. F. Romero

Introdução

- Tarefa de navegação de um robô móvel
  - Competência fundamental de se movimentar por um ambiente
  - Evitar colisões
  - Atingir localidades específicas
  - Atuadores e sensores

Técnicas de localização

- Como o robô sabe onde se encontra à medida que se movimenta em um determinado ambiente?
  - Executa ciclos que intercalam movimento com percepção



Técnicas de localização

- Técnicas probabilísticas permitem explicitar as probabilidades de cada possível localização

- Estimação Bayesiana, Filtro de Kalman e Localização de Monte Carlo

- O Filtro de Kalman e sua versão estendida estimam o estado de uma variável aleatória dinâmica, cuja densidade de probabilidade é gaussiana, a partir da sequência de observações ruidosas

- No caso da localização de um robô, o intuito é estimar a sua localização a cada instante, X_t, que por sua vez influencia diretamente na observação dos sensores naquele momento, Z_t

Filtro de Kalman

- Problema de filtragem
  - Dado um modelo dinâmico estocástico x_k, a sequência de observações ruidosas Z_k e as condições iniciais, encontrar $\hat{x}_k$, o valor estimado x_k
- Considere o seguinte sistema não-linear, descrito pela equação de diferenças e o modelo de observação com ruído:

$$\begin{aligned} \mathbf{x}_k &= \mathbf{M}(\mathbf{x}_{k-1}) + \mathbf{w}_k \\ \mathbf{z}_k &= \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \end{aligned}$$

- O estado inicial x_0 é um vetor aleatório com média $m_0 = E[x_0]$ e covariância $E[(x_0-m_0)(x_0-m_0)^T]$

Filtro de Kalman

- w_k captura incertezas no modelo e v_k denota o ruído da medida
- w_k e v_k não são correlacionados:

$$E[\mathbf{w}_k] = 0 \quad E[\mathbf{w}_k\mathbf{w}_k^T] = \mathbf{Q}_k \quad E[\mathbf{w}_k\mathbf{w}_j^T] = 0 \text{ for } k \neq j \quad E[\mathbf{w}_k\mathbf{x}_0^T] = 0 \text{ for all } k$$

$$E[\mathbf{v}_k] = 0 \quad E[\mathbf{v}_k\mathbf{v}_k^T] = \mathbf{R}_k \quad E[\mathbf{v}_k\mathbf{v}_j^T] = 0 \text{ for } k \neq j \quad E[\mathbf{v}_k\mathbf{x}_0^T] = 0 \text{ for all } k$$

- Também os vetores aleatórios w_k e v_j^T não são correlacionados:

$$E[\mathbf{w}_k\mathbf{v}_j^T] = 0 \text{ for all } k \text{ and } j$$

Filtro de Kalman

- Dimensão e descrição das variáveis:

x_k	$n \times 1$	Vetor de estados
w_k	$n \times 1$	Vetor de ruído do processo
z_k	$m \times 1$	Vetor de observação
v_k	$m \times 1$	Vetor de ruído da medida
$M(.)$	$n \times 1$	Função vetorial do processo não-linear
$H(.)$	$m \times 1$	Função vetorial da observação não-linear
Q_k	$n \times n$	Matriz de covariância de ruído do processo
R_k	$m \times m$	Matriz de covariância de medida do processo

Derivação EKF

- Assumindo que as não-linearidades na dinâmica e no modelo de observação são suaves, M(x_k) e h(x_k) serão expandidas em Séries de Taylor e, assim, aproximadas a previsão e a próxima estimativa de x_k
- Modelo do passo de previsão
  - Informação disponível inicialmente:

$$\begin{aligned}\hat{\mathbf{x}}_0 &= \mathbf{m}_0 = E[\mathbf{x}_0] \\ \mathbf{P}_0 &= E[(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T]\end{aligned}$$

Modelo do passo de previsão

- Assumindo que a estimativa ótima $\hat{\mathbf{x}}_{k-1} \equiv E[\mathbf{x}_{k-1}|\mathbf{Z}_{k-1}]$ com covariância P_{k-1}, no tempo $k-1$. A parte previsível de x_k é dada por:

$$
\begin{aligned}
\mathbf{x}_k^f &\equiv E[\mathbf{x}_k|\mathbf{Z}_{k-1}] \\
&= E[\mathbf{M}(\mathbf{x}_{k-1}) + \mathbf{w}_k|\mathbf{Z}_{k-1}] \\
&= E[\mathbf{M}(\mathbf{x}_{k-1})|\mathbf{Z}_{k-1}]
\end{aligned}
$$

- Expandindo M(.) em Séries de Taylor sobre $\hat{\mathbf{x}}_{k-1}$:

$$\mathbf{M}(\mathbf{x}_{k-1}) \equiv \mathbf{M}(\hat{\mathbf{x}}_{k-1}) + \mathbf{J}_{\mathbf{M}}(\hat{\mathbf{x}}_{k-1})(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}) + \mathbf{H.O.T.}$$

- J_M é o Jacobiano de M(.)
- Os termos de ordem mais alta são desprezados

Modelo do passo de previsão

- O Jacobiano é definido por:

$$\mathbf{J_M} \equiv \begin{bmatrix} \frac{\partial \mathbf{M}_1}{\partial x_1} & \frac{\partial \mathbf{M}_1}{\partial x_2} & \cdots & \frac{\partial \mathbf{M}_1}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial \mathbf{M}_n}{\partial x_1} & \frac{\partial \mathbf{M}_n}{\partial x_2} & \cdots & \frac{\partial \mathbf{M}_n}{\partial x_n} \end{bmatrix}$$

$$\mathbf{M}(\mathbf{x}) = (\mathbf{M}_1(\mathbf{x}), \mathbf{M}_2(\mathbf{x}), \ldots, \mathbf{M}_n(\mathbf{x}))^T$$

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$$

$$\mathbf{M}(\mathbf{x}_{k-1}) \approx \mathbf{M}(\hat{\mathbf{x}}_{k-1}) + \mathbf{J_M}(\hat{\mathbf{x}}_{k-1})\mathbf{e}_{k-1}$$

Modelo do passo de previsão

- Tomando a esperança em ambos os lados, condicionada por Z_{k-1}:

$$E[\mathbf{M}(\mathbf{x}_{k-1})|\mathbf{Z}_{k-1}] \approx \mathbf{M}(\hat{\mathbf{x}}_{k-1}) + \mathbf{J}_{\mathbf{M}}(\hat{\mathbf{x}}_{k-1})E[\mathbf{e}_{k-1}|\mathbf{Z}_{k-1}]$$

em que $E[e_{k-1}|Z_{k-1}] = 0$

- Assim, o valor previsto de x_k é:

$$\mathbf{x}_k^f \approx \mathbf{M}(\hat{\mathbf{x}}_{k-1})$$

- E o erro predito fica:

$$\begin{aligned} \mathbf{e}_k^f &\equiv \mathbf{x}_k - \mathbf{x}_k^f \\ &= \mathbf{M}(\mathbf{x}_{k-1}) + \mathbf{w}_k - \mathbf{M}(\hat{\mathbf{x}}_{k-1}) \\ &\approx \mathbf{J}_{\mathbf{M}}(\hat{\mathbf{x}}_{k-1})\mathbf{e}_{k-1} + \mathbf{w}_k \end{aligned}$$

Modelo do passo de previsão

- A covariância do erro prevista:

$$
\begin{aligned}
\mathbf{P}_k^f &\equiv E[\mathbf{e}_k^f(\mathbf{e}_k^f)^T] \\
&= \mathbf{J}_{\mathbf{M}}(\hat{\mathbf{x}}_{k-1})E[\mathbf{e}_{k-1}\mathbf{e}_{k-1}^T]\mathbf{J}_{\mathbf{M}}^T(\hat{\mathbf{x}}_{k-1}) + E[\mathbf{w}_k\mathbf{w}_k^T] \\
&= \mathbf{J}_{\mathbf{M}}(\hat{\mathbf{x}}_{k-1})\mathbf{P}_{k-1}\mathbf{J}_{\mathbf{M}}^T(\hat{\mathbf{x}}_{k-1}) + \mathbf{Q}_k
\end{aligned}
$$

Passo de assimilação de dados

- No tempo *k*, temos duas partes de informação:
  - O valor previsto $\mathbf{x}_k^f$ com covariância $\mathbf{P}_k^f$
  - Medida z_k com covariância R_k
- O objetivo é encontrar a melhor estimativa não tendenciosa, no sentido dos mínimos quadrados $\hat{\mathbf{x}}_k$ de $\mathbf{x}_k$
- Uma maneira é assumir que a estimativa é uma combinação linear de $\mathbf{x}_k^f$ e $\mathbf{z}_k$

Passo de assimilação de dados

- Seja:

$$\hat{\mathbf{x}}_k = \mathbf{a} + \mathbf{K}_k \mathbf{z}_k$$

- Considerando que a estimativa não é tendenciosa:

$$\begin{aligned} 0 &= E[\mathbf{x}_k - \hat{\mathbf{x}}_k | \mathbf{Z}_k] \\ &= E[(\mathbf{x}_k^f + \mathbf{e}_k^f) - (\mathbf{a} + \mathbf{K}_k \mathbf{h}(\mathbf{x}_k) + \mathbf{K}_k \mathbf{v}_k) | \mathbf{Z}_k] \\ &= \mathbf{x}_k^f - \mathbf{a} - \mathbf{K}_k E[\mathbf{h}(\mathbf{x}_k) | \mathbf{Z}_k] \\ \mathbf{a} &= \mathbf{x}_k^f - \mathbf{K}_k E[\mathbf{h}(\mathbf{x}_k) | \mathbf{Z}_k] \end{aligned}$$

- Substituindo o valor de a:

$$\hat{\mathbf{x}}_k = \mathbf{x}_k^f + \mathbf{K}_k (\mathbf{z}_k - E[\mathbf{h}(\mathbf{x}_k) | \mathbf{Z}_k])$$

Passo de assimilação de dados

- Expandindo h(.) em Séries de Taylor sobre $\mathbf{x}_k^f$

$$\mathbf{h}(\mathbf{x}_k) \equiv \mathbf{h}(\mathbf{x}_k^f) + \mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f)(\mathbf{x}_k - \mathbf{x}_k^f) + \mathrm{H.O.T.}$$

- O Jacobiano de h(.) é definido como:

$$\mathbf{J}_{\mathbf{h}} \equiv \begin{bmatrix} \frac{\partial \mathbf{h}_1}{\partial x_1} & \frac{\partial \mathbf{h}_1}{\partial x_2} & \cdots & \frac{\partial \mathbf{h}_1}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial \mathbf{h}_m}{\partial x_1} & \frac{\partial \mathbf{h}_m}{\partial x_2} & \cdots & \frac{\partial \mathbf{h}_m}{\partial x_n} \end{bmatrix}$$

$$\mathbf{h}(\mathbf{x}) = (\mathbf{h}_1(\mathbf{x}), \mathbf{h}_2(\mathbf{x}), \ldots, \mathbf{h}_m(\mathbf{x}))^T$$

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$$

Passo de assimilação de dados

- Tomando a esperança dos dois lados da equação de $\mathbf{h}(\mathbf{x}_k)$ condicionada por Z_k:

$$E[\mathbf{h}(\mathbf{x}_k)|\mathbf{Z}_k] \approx \mathbf{h}(\mathbf{x}_k^f) + \mathbf{J_h}(\mathbf{x}_k^f)E[\mathbf{e}_k^f|\mathbf{Z}_k]$$

- em que $E[\mathbf{e}_k^f|\mathbf{Z}_k] = 0$.
- Substituindo em $\hat{\mathbf{x}}_k = \mathbf{x}_k^f + \mathbf{K}_k(\mathbf{z}_k - E[\mathbf{h}(\mathbf{x}_k)|\mathbf{Z}_k])$, o estado estimado é:

$$\hat{\mathbf{x}}_k \quad \approx \quad \mathbf{x}_k^f + \mathbf{K}_k(\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k^f))$$

Passo de assimilação de dados

- O erro no valor estimado de $\hat{\mathbf{x}}_k$ é:

$$
\begin{aligned}
\mathbf{e}_k &\equiv \mathbf{x}_k - \hat{\mathbf{x}}_k \\
&= \mathbf{M}(\mathbf{x}_{k-1}) + \mathbf{w}_k - \mathbf{x}_k^f - \mathbf{K}_k(\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k^f)) \\
&\approx \mathbf{M}(\mathbf{x}_{k-1}) - \mathbf{M}(\hat{\mathbf{x}}_{k-1}) + \mathbf{w}_k - \mathbf{K}_k(\mathbf{h}(\mathbf{x}_k) - \mathbf{h}(\mathbf{x}_k^f) + \mathbf{v}_k) \\
&\approx \mathbf{J}_{\mathbf{M}}(\hat{\mathbf{x}}_{k-1})\mathbf{e}_{k-1} + \mathbf{w}_k - \mathbf{K}_k(\mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f)\mathbf{e}_k^f + \mathbf{v}_k) \\
&\approx \mathbf{J}_{\mathbf{M}}(\hat{\mathbf{x}}_{k-1})\mathbf{e}_{k-1} + \mathbf{w}_k - \mathbf{K}_k\mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f)(\mathbf{J}_{\mathbf{M}}(\hat{\mathbf{x}}_{k-1})\mathbf{e}_{k-1} + \mathbf{w}_k) - \mathbf{K}_k\mathbf{v}_k \\
&\approx (\mathbf{I} - \mathbf{K}_k\mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f))\mathbf{J}_{\mathbf{M}}(\hat{\mathbf{x}}_{k-1})\mathbf{e}_{k-1} + (\mathbf{I} - \mathbf{K}_k\mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f))\mathbf{w}_k - \mathbf{K}_k\mathbf{v}_k
\end{aligned}
$$

Passo de assimilação de dados

- Então, a covariância posterior da nova estimativa é:

$$
\begin{aligned}
\mathbf{P}_k \equiv\ & E[\mathbf{e}_k\mathbf{e}_k^T] \\
=\ & (\mathbf{I}-\mathbf{K}_k\mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f))\mathbf{J}_{\mathrm{M}}(\hat{\mathbf{x}}_{k-1})\mathbf{P}_{k-1}\mathbf{J}_{\mathrm{M}}^T(\hat{\mathbf{x}}_{k-1})(\mathbf{I}-\mathbf{K}_k\mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f))^T \\
& +(\mathbf{I}-\mathbf{K}_k\mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f))\mathbf{Q}_k(\mathbf{I}-\mathbf{K}_k\mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f))^T+\mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \\
=\ & (\mathbf{I}-\mathbf{K}_k\mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f))\mathbf{P}_k^f(\mathbf{I}-\mathbf{K}_k\mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f))^T+\mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \\
=\ & \mathbf{P}_k^f-\mathbf{K}_k\mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f)\mathbf{P}_k^f-\mathbf{P}_k^f\mathbf{J}_{\mathbf{h}}^T(\mathbf{x}_k^f)\mathbf{K}_k^T+\mathbf{K}_k\mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f)\mathbf{P}_k^f\mathbf{J}_{\mathbf{h}}^T(\mathbf{x}_k^f)\mathbf{K}_k^T+\mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T
\end{aligned}
$$

- Essa fórmula é válida para qualquer K_k
- Como no filtro de Kalman padrão, K_k é descoberto por minimizar o *tr*(P_k)

$$
\begin{aligned}
0 &= \frac{\partial tr(\mathbf{P}_k)}{\partial \mathbf{K}_k} \\
&= -(\mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f)\mathbf{P}_k^f)^T-\mathbf{P}_k^f\mathbf{J}_{\mathbf{h}}^T(\mathbf{x}_k^f)+2\mathbf{K}_k\mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f)\mathbf{P}_k^f\mathbf{J}_{\mathbf{h}}^T(\mathbf{x}_k^f)+2\mathbf{K}_k\mathbf{R}_k
\end{aligned}
$$

Passo de assimilação de dados

- Substituindo novamente em P_k, tem-se:

$$
\begin{aligned}
\mathbf{P}_k &= (\mathbf{I} - \mathbf{K}_k \mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f))\mathbf{P}_k^f - (\mathbf{I} - \mathbf{K}_k \mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f))\mathbf{P}_k^f \mathbf{J}_{\mathbf{h}}^T(\mathbf{x}_k^f)\mathbf{K}_k^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \\
&= (\mathbf{I} - \mathbf{K}_k \mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f))\mathbf{P}_k^f - \left(\mathbf{P}_k^f \mathbf{J}_{\mathbf{h}}^T(\mathbf{x}_k^f) - \mathbf{K}_k \mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f)\mathbf{P}_k^f \mathbf{J}_{\mathbf{h}}^T(\mathbf{x}_k^f) - \mathbf{K}_k \mathbf{R}_k\right)\mathbf{K}_k^T \\
&= (\mathbf{I} - \mathbf{K}_k \mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f))\mathbf{P}_k^f - \left[\mathbf{P}_k^f \mathbf{J}_{\mathbf{h}}^T(\mathbf{x}_k^f) - \mathbf{K}_k \left(\mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f)\mathbf{P}_k^f \mathbf{J}_{\mathbf{h}}^T(\mathbf{x}_k^f) + \mathbf{R}_k\right)\right]\mathbf{K}_k^T \\
&= (\mathbf{I} - \mathbf{K}_k \mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f))\mathbf{P}_k^f - \left[\mathbf{P}_k^f \mathbf{J}_{\mathbf{h}}^T(\mathbf{x}_k^f) - \mathbf{P}_k^f \mathbf{J}_{\mathbf{h}}^T(\mathbf{x}_k^f)\right]\mathbf{K}_k^T \\
&= (\mathbf{I} - \mathbf{K}_k \mathbf{J}_{\mathbf{h}}(\mathbf{x}_k^f))\mathbf{P}_k^f
\end{aligned}
$$