Atrasos em filtros de Kalman

Gustavo T. Pfeiffer

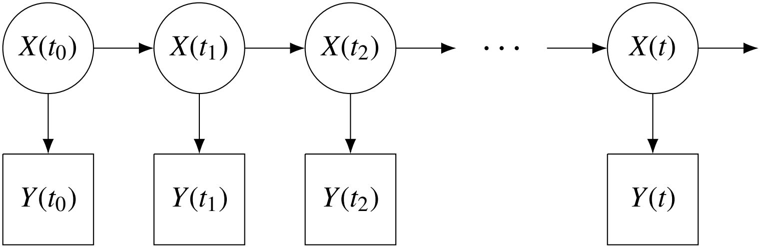
19/8/2025

Agenda

- 1. Filtro de Kalman: O que é, para que serve
- 2. Modelo matemático
- 3. Problema dos atrasos
- 4. Solução

O que é um filtro de Kalman

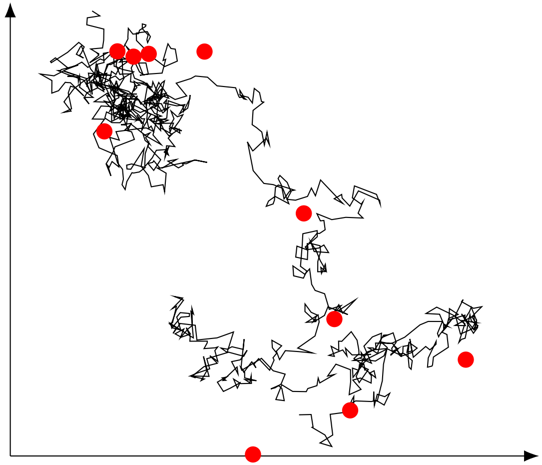
- Sistema $X(t)$ evolui estocasticamente no tempo



- Medidas $Y(t)$ com erros aleatórios
- Objetivo: Estimar $X(t)$ dadas todas as medições $Y(0)$, ..., $Y(t)$ em tempo real

Exemplo

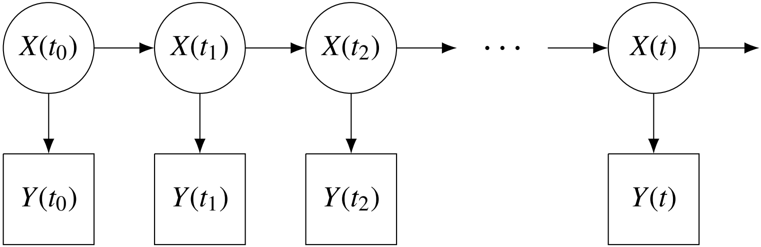
Movimento browniano



Aplicações

- Tracking de objetos com câmera, LiDAR, etc.
- Posicionamento (GPS, acelerômetro, etc.)
- Sistemas complexos (ex.: reservatórios de petróleo)

Hipóteses



- Hipóteses
  - Erros de medição são independentes entre si
  - Todas as variáveis são gaussianas (normais multivariadas)
  - Relações entre as variáveis são lineares
- Modelos não-lineares são muito usados na prática, porém são aproximações

Modelo matemático

Modelo matemático

- Evolução no tempo
    - $X(t + \Delta t) = TX(t) + \mathcal{N}(0, A)$
- Observações
    - $Y(t) = PX(t) + \mathcal{N}(0, B)$

Modelo matemático

- Evolução no tempo
  - $X(t + \Delta t) = TX(t) + \mathcal{N}(0, A)$
- Observações
  - $Y(t) = PX(t) + \mathcal{N}(0, B)$
- Exemplo:
  - Partícula com posição s e velocidade v
  - Velocidade v muda aleatoriamente (movimento browniano)
    - Posição s tem movimento mais suave (C_1)
  - Apenas a posição s é medida

Modelo matemático

- Evolução no tempo
  - $X(t+\Delta t) = TX(t) + \mathcal{N}(0, A)$
- Observações
  - $Y(t) = PX(t) + \mathcal{N}(0, B)$
- Exemplo:
  - Partícula com posição s e velocidade v
  - Velocidade v muda aleatoriamente (movimento browniano)
    - Posição s tem movimento mais suave (C_1)
  - Apenas a posição s é medida
  - Modelo resultante:
    - $X(t) = \begin{bmatrix} v(t) \\ s(t) \end{bmatrix}$, $T = \begin{bmatrix} 1 & 0 \\ \Delta t & 1 \end{bmatrix}$, $P = \begin{bmatrix} 0 & 1 \end{bmatrix}$
    - $A = \begin{bmatrix} \sigma^2 \Delta t & \sigma^2 \Delta t^2/2 \\ \sigma^2 \Delta t^2/2 & \sigma^2 \Delta t^3/3 \end{bmatrix}$

Inferência

- Estimativa da distribuição de $X(t)$ é representada por uma média μ (vetor) e variância Σ (matriz)
- Na passagem do tempo sem observações, a atualização *a priori* (ou predição) é trivial
  - $\mathrm{E}[X(t+\Delta t)] = T\mathrm{E}[X(t)]$
  - $\mathrm{Var}[X(t+\Delta t)] = T\mathrm{Var}[X(t)]T^T + A$
- Ao chegar uma nova observação $Y(t)$, a atualização *a posteriori* é feita pela lei de Bayes assumindo distribuição gaussiana
  - $\mathrm{Var}[X(t)|Y(t)]^{-1} = \mathrm{Var}[X(t)]^{-1} + P^T B^{-1} P$
  - $\mathrm{E}[X(t)|Y(t)] = \mathrm{Var}[X(t)|Y(t)]\left(\mathrm{Var}[X(t)]^{-1}\mathrm{E}[X(t)] + P^T B^{-1} Y(t)\right)$

Inferência

- Estimativa da distribuição de $X(t)$ é representada por uma média μ (vetor) e variância Σ (matriz)
- Na passagem do tempo sem observações, a atualização *a priori* (ou predição) é trivial
  - $\mathrm{E}[X(t+\Delta t)] = T\mathrm{E}[X(t)]$
  - $\mathrm{Var}[X(t+\Delta t)] = T\mathrm{Var}[X(t)]T^T + A$
- Ao chegar uma nova observação $Y(t)$, a atualização *a posteriori* é feita pela lei de Bayes assumindo distribuição gaussiana
  - $\mathrm{Var}[X(t)|Y(t)]^{-1} = \mathrm{Var}[X(t)]^{-1} + P^T B^{-1} P$
  - $\mathrm{E}[X(t)|Y(t)] = \mathrm{Var}[X(t)|Y(t)] \left(\mathrm{Var}[X(t)]^{-1}\mathrm{E}[X(t)] + P^T B^{-1} Y(t)\right)$
- Contemple a dualidade entre esses dois tipos de eventos!
  - Um (*a priori*) atualiza a variância linearmente
  - O outro (*a posteriori*) atualiza a inversa linearmente

Um parêntese

Um parêntese

- Situação I (adição de erro):

$$X + Y,\ X \perp\!\!\!\perp Y$$

$$\mathrm{E}[X+Y] = \mathrm{E}[X] + \mathrm{E}[Y]$$

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$$

Um parêntese

- Situação I (adição de erro):

$$X + Y,\ X \perp\!\!\!\perp Y$$

$$\mathrm{E}[X + Y] = \mathrm{E}[X] + \mathrm{E}[Y]$$

$$\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$$

- Situação II (adição de informação):

$$X = \mathcal{N}(z, \Sigma_X),\ Y = \mathcal{N}(z, \Sigma_Y),\ X \perp\!\!\!\perp Y$$

$$\widehat{z} = \arg\max_z P(X, Y; z)$$

$$\mathrm{Var}[\widehat{z}]^{-1} = \Sigma_X^{-1} + \Sigma_Y^{-1}$$

$$\mathrm{Var}[\widehat{z}]^{-1}\widehat{z} = \Sigma_X^{-1} X + \Sigma_Y^{-1} Y$$

O problema dos atrasos

O problema dos atrasos

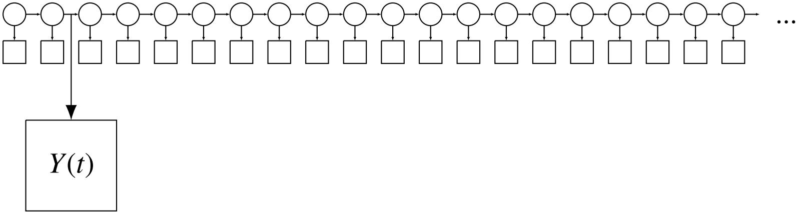
- Na prática, observações podem chegar com atraso e fora de ordem
  - Atrasos de rede, processamento, etc.

O problema dos atrasos

- Na prática, observações podem chegar com atraso e fora de ordem
  - Atrasos de rede, processamento, etc.
- Questionamentos
  - É possível computar eventos de observação fora de ordem?
  - Se atrasos forem frequentes e substanciais, qual é a forma mais eficiente de manter (μ, Σ) atualizada?
  - Em caso de um atraso muito longo, é possível atualizar (μ, Σ) sem ter que recalcular tudo?

O problema dos atrasos

- Na prática, observações podem chegar com atraso e fora de ordem
  - Atrasos de rede, processamento, etc.
- Questionamentos
  - É possível computar eventos de observação fora de ordem?
  - Se atrasos forem frequentes e substanciais, qual é a forma mais eficiente de manter (μ, Σ) atualizada?
  - Em caso de um atraso muito longo, é possível atualizar (μ, Σ) sem ter que recalcular tudo?



Meta-modelo

- Como modelar uma sequência de eventos de passagem do tempo e observação em uma única operação?
  - Passagem do tempo (*a priori*):

$$\Sigma_{\text{new}} = T\Sigma_{\text{old}}T^T + A$$

$$\mu_{\text{new}} = T\mu_{\text{old}}$$

  - Observação (*a posteriori*):

$$\Sigma_{\text{new}}^{-1} = \Sigma_{\text{old}}^{-1} + P^T B^{-1} P$$

$$\mu_{\text{new}} = \Sigma_{\text{new}}(\Sigma_{\text{old}}^{-1} + P^T B^{-1} Y)$$

Meta-modelo

- Como modelar uma sequência de eventos de passagem do tempo e observação em uma única operação?
  - Passagem do tempo (*a priori*):

$$\Sigma_{\text{new}} = T\Sigma_{\text{old}}T^T + A$$

$$\mu_{\text{new}} = T\mu_{\text{old}}$$

  - Observação (*a posteriori*):

$$\Sigma_{\text{new}}^{-1} = \Sigma_{\text{old}}^{-1} + P^T B^{-1} P$$

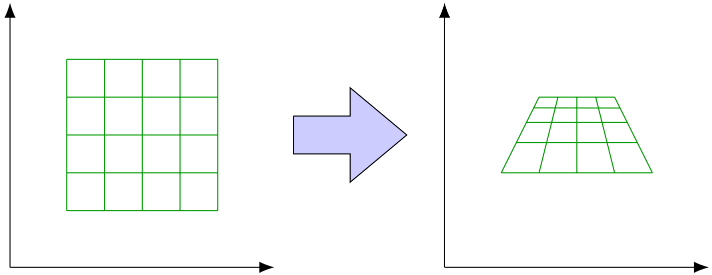
$$\mu_{\text{new}} = \Sigma_{\text{new}}(\Sigma_{\text{old}}^{-1} + P^T B^{-1} Y)$$

- Vamos pegar emprestado um truque da área de processamento de imagens

Coordenadas homogêneas

- Adicionar uma dimensão para permitir transformações projetivas (não-lineares)

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \simeq \begin{bmatrix} ax \\ ay \\ a \end{bmatrix} \quad (a \neq 0)$$



A ideia

- Decompor $\Sigma = U^{-1}V$ e só atualizar U ou V, conforme necessário
  - A redundância é intencional
- Semelhantemente, decompor $\mu = U^{-1}z$

A ideia

- Decompor $\Sigma = U^{-1}V$ e só atualizar U ou V, conforme necessário
  - A redundância é intencional
- Semelhantemente, decompor $\mu = U^{-1}z$
- Resultado: Os eventos viram multiplicações de matrizes
  - *A priori*:

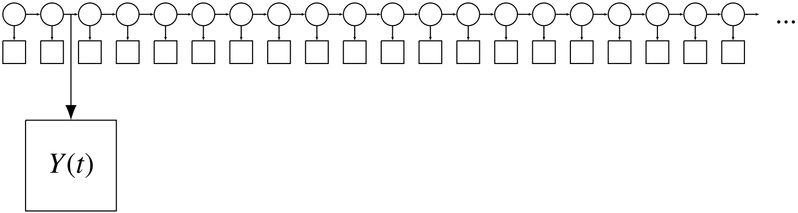
$$\begin{bmatrix} U_{\text{new}} & V_{\text{new}} & z_{\text{new}} \end{bmatrix} = \begin{bmatrix} U_{\text{old}} & V_{\text{old}} & z_{\text{old}} \end{bmatrix} \begin{bmatrix} T^{-1} & T^{-1}A & 0 \\ 0 & T^T & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

  - *A posteriori*:

$$\begin{bmatrix} U_{\text{new}} & V_{\text{new}} & z_{\text{new}} \end{bmatrix} = \begin{bmatrix} U_{\text{old}} & V_{\text{old}} & z_{\text{old}} \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ P^T B^{-1} P & I & P^T B^{-1} Y \\ 0 & 0 & 1 \end{bmatrix}$$

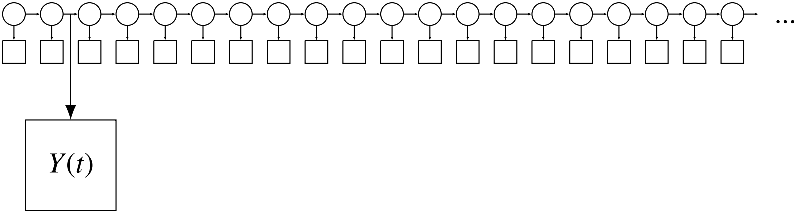
Conclusão

- Uma sequência de atualizações *a priori* e *a posteriori* é um produtório de matrizes
- Multiplicação de matrizes não é comutativa (a ordem importa!)
- Logo, não é possível manter essas atualizações em $O(1)$



Conclusão

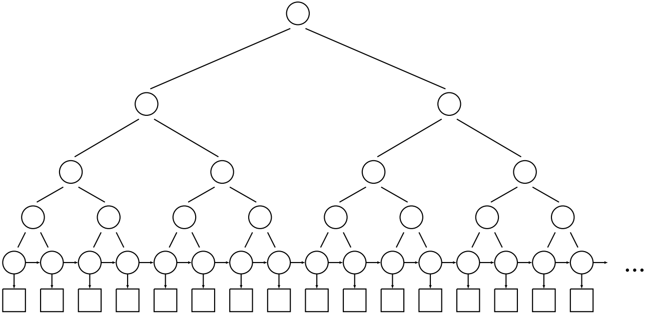
- Uma sequência de atualizações *a priori* e *a posteriori* é um produtório de matrizes
- Multiplicação de matrizes não é comutativa (a ordem importa!)
- Logo, não é possível manter essas atualizações em $O(1)$



- No entanto, é possível atualizar em $O(\log(n))$. Como?

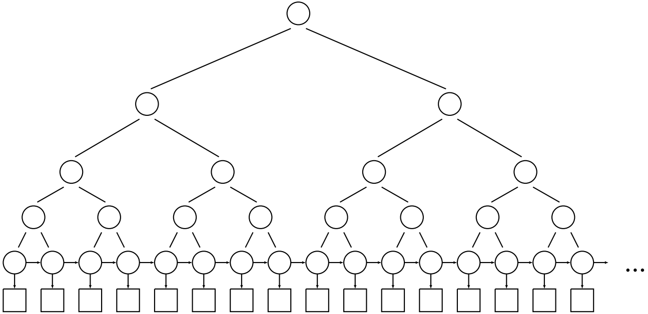
Conclusão

Usando uma árvore!



Conclusão

Usando uma árvore!



- Que tipo de árvore?

Considerações

Considerações

- Por que essa ideia não é usada na prática?
  - Alto custo de memória
  - Modelo precisa ser gaussiano e linear

Considerações

- Por que essa ideia não é usada na prática?
    - Alto custo de memória
    - Modelo precisa ser gaussiano e linear
- Ideia não é inédita na literatura
    - Representações das atualizações do filtro de Kalman por meio de um "operador associativo" já foram propostas
    - Aplicações incluem paralelização do filtro de Kalman em GPU (ex.: Särkkä e García-Fernández, 2020)