



Redes de Sensores

Filtrado no lineal

Manuel A. Vázquez
Joaquín Míguez
Jose Miguel Leiva

4 de febrero de 2024

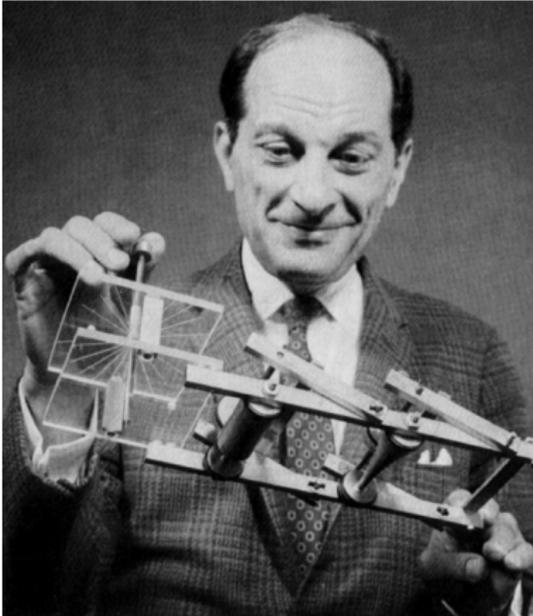
Índice

- 1 Un mundo lineal
- 2 Filtro de Kalman extendido
- 3 Filtro de Kalman unscented
- 4 Un planteamiento más general del problema de estimación
- 5 Monte Carlo
- 6 Importance sampling
- 7 Filtrado de partículas

Índice

- 1 Un mundo lineal
- 2 Filtro de Kalman extendido
- 3 Filtro de Kalman unscented
- 4 Un planteamiento más general del problema de estimación
- 5 Monte Carlo
- 6 Importance sampling
- 7 Filtrado de partículas

Linealidad



“Utilizar un término como ciencia no lineal es como referirse al grueso de la zoología como el estudio de los animales que no son elefantes.”

— Stanislaw Ulam

Índice

- 1 Un mundo lineal
- 2 Filtro de Kalman extendido**
- 3 Filtro de Kalman unscented
- 4 Un planteamiento más general del problema de estimación
- 5 Monte Carlo
- 6 Importance sampling
- 7 Filtrado de partículas

Modelo dinámico no lineal

Consideramos la misma ecuación de estado que antes

- $$\mathbf{x}_t = \mathbf{F}\mathbf{x}_{t-1} + \mathbf{v}_t,$$

...pero ahora la relación entre el estado y las observaciones viene dada por la función (vectorial) $\mathbf{h} : \mathbb{R}^M \rightarrow \mathbb{R}^N$ (más ruido gaussiano aditivo)

- $$\mathbf{y}_t = \mathbf{h}(\mathbf{x}_t) + \mathbf{w}_t,$$

siendo \mathbf{h} un vector de funciones escalares de variable vectorial,

$$\mathbf{h}(\mathbf{x}_t) = \begin{bmatrix} h_1(\mathbf{x}_t) \\ h_2(\mathbf{x}_t) \\ \vdots \\ h_N(\mathbf{x}_t) \end{bmatrix}$$

No podemos aplicar el filtro de Kalman!!

Modelo dinámico linealizado

Objetivo

Aplicar el KF sobre el modelo no lineal para estimar \mathbf{x}_t dado $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_t$

Podemos construir una aproximación lineal de la ecuación de observación¹ utilizando una **serie de Taylor** de primer orden

$$\mathbf{h}(\mathbf{x}_t) \approx \mathbf{h}(\mathbf{x}^0) + \left[\frac{\partial \mathbf{h}}{\partial \mathbf{x}_t} \right]_{\mathbf{x}_t = \mathbf{x}^0} (\mathbf{x}_t - \mathbf{x}^0),$$

siendo

$$\left[\frac{\partial \mathbf{h}}{\partial \mathbf{x}_t} \right] = \begin{bmatrix} \frac{\partial h_1}{\partial x_{1,t}} & \frac{\partial h_1}{\partial x_{2,t}} & \dots & \frac{\partial h_1}{\partial x_{M,t}} \\ \frac{\partial h_2}{\partial x_{1,t}} & \frac{\partial h_2}{\partial x_{2,t}} & \dots & \frac{\partial h_2}{\partial x_{M,t}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_N}{\partial x_{1,t}} & \frac{\partial h_N}{\partial x_{2,t}} & \dots & \frac{\partial h_N}{\partial x_{M,t}} \end{bmatrix}$$

es el Jacobiano (de derivadas parciales) de \mathbf{h} .

¹Podríamos hacer lo mismo con una ecuación de estado no lineal!!!

Derivando el filtro de Kalman extendido

EKF define unas observaciones transformadas,

$$\tilde{\mathbf{y}}_t = \mathbf{y}_t - \mathbf{h}(\mathbf{x}^0) + \left[\frac{\partial \mathbf{h}}{\partial \mathbf{x}_t} \right]_{\mathbf{x}_t = \mathbf{x}^0} \mathbf{x}^0,$$

que dan lugar a un modelo dinámico aproximado que es lineal y gaussiano

$$\mathbf{x}_t = \mathbf{F}\mathbf{x}_{t-1} + \mathbf{v}_t$$

$$\tilde{\mathbf{y}}_t = \left[\frac{\partial \mathbf{h}}{\partial \mathbf{x}_t} \right]_{\mathbf{x}_t = \mathbf{x}^0} \mathbf{x}_t + \mathbf{w}_t$$



Conseguido!!

Es inmediato aplicar el KF sobre el modelo anterior.

Filtro de Kalman extendido

- Predicción

$$\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}\hat{\mathbf{x}}_{t-1|t-1}$$

$$\mathbf{P}_{t|t-1} = \mathbf{Q} + \mathbf{F}\mathbf{P}_{t-1|t-1}\mathbf{F}^\top$$

- Actualización

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} \left[\frac{\partial \mathbf{h}}{\partial \mathbf{x}_t} \right]_{\mathbf{x}_t=\mathbf{x}^0}^\top \left(\left[\frac{\partial \mathbf{h}}{\partial \mathbf{x}_t} \right]_{\mathbf{x}_t=\mathbf{x}^0} \mathbf{P}_{t|t-1} \left[\frac{\partial \mathbf{h}}{\partial \mathbf{x}_t} \right]_{\mathbf{x}_t=\mathbf{x}^0}^\top + \mathbf{R} \right)^{-1}$$

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t(\mathbf{y}_t - \mathbf{h}(\hat{\mathbf{x}}_{t|t-1}))$$

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{K}_t \left(\left[\frac{\partial \mathbf{h}}{\partial \mathbf{x}_t} \right]_{\mathbf{x}_t=\mathbf{x}^0} \mathbf{P}_{t|t-1} \left[\frac{\partial \mathbf{h}}{\partial \mathbf{x}_t} \right]_{\mathbf{x}_t=\mathbf{x}^0}^\top \right) \mathbf{K}_n^\top,$$

siendo \mathbf{Q} la covarianza de \mathbf{v}_t y \mathbf{R} la de \mathbf{w}_t .

- El punto de linealización \mathbf{x}^0 tiene que ser suficientemente cercano a \mathbf{x}_t para que el algoritmo funcione correctamente. Normalmente es suficiente tomar $\mathbf{x}^0 = \hat{\mathbf{x}}_{t|t-1}$.

Ejemplo de función no lineal en seguimiento

Objetivo

Localización y seguimiento de un objeto con velocidad constante y conocida \mathbf{c} .

 S_2 ×

 S_4
 S_1
 S_3

La misma ecuación de estado que antes,

- $$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{c}T + \mathbf{v}_t,$$

...una ecuación de observación más *realista* basada en Received Signal Strength Indicator (RSSI),

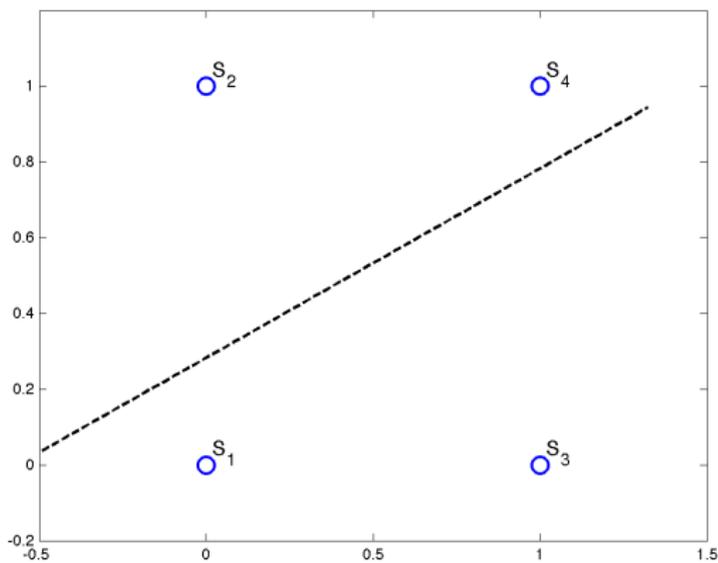
- $$y_{t,i} = \underbrace{k_1 - k_2 \log \|\mathbf{x}_t - \mathbf{s}_i\|}_{\text{RSSI}_i} + w_{t,i}, \quad i = 1, \dots, N$$

con k_1 y k_2 constantes conocidas y \mathbf{s}_i la posición del sensor correspondiente.

(antes, $\mathbf{y}_t = \mathbf{x}_t + \mathbf{w}_t$)

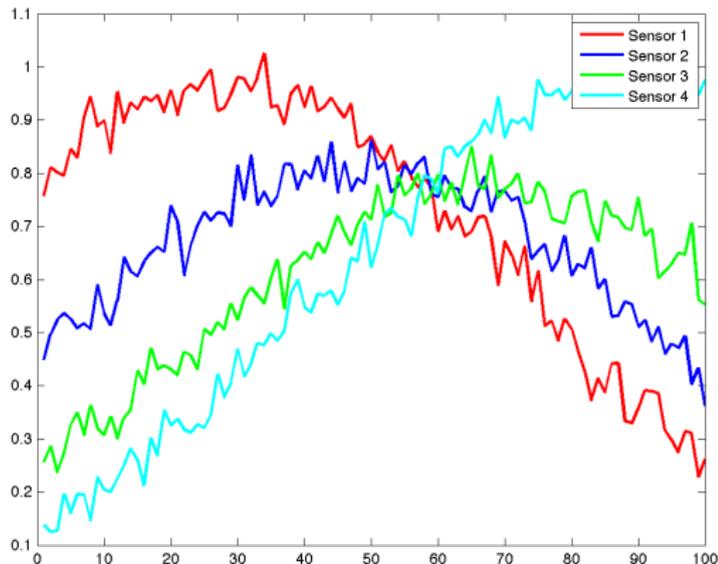
Ejemplo

- Trayectoria real



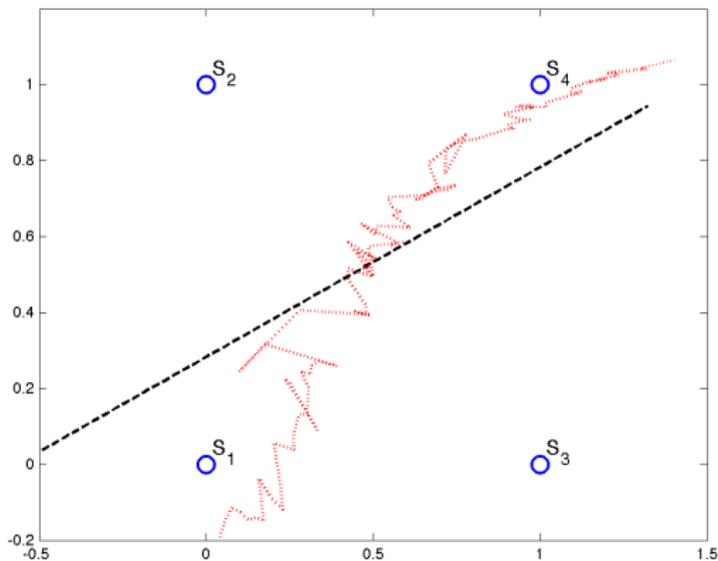
Ejemplo

- Lectura de los sensores



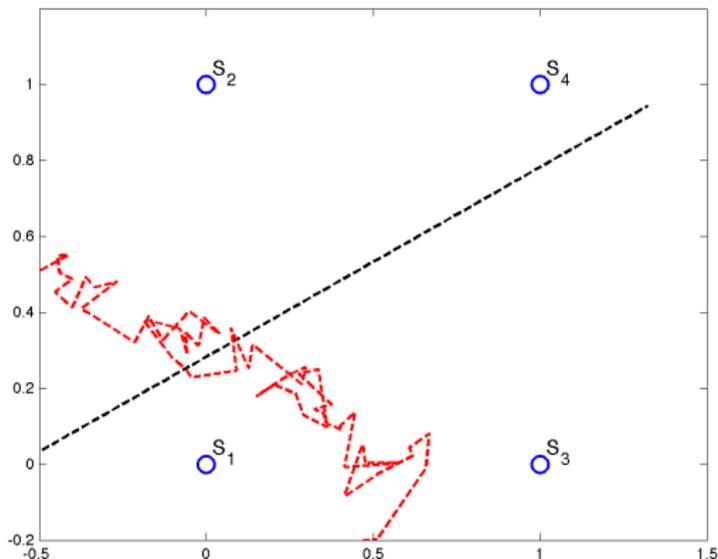
Ejemplo

- Resultado de filtrado usando solo el Sensor 2



Ejemplo

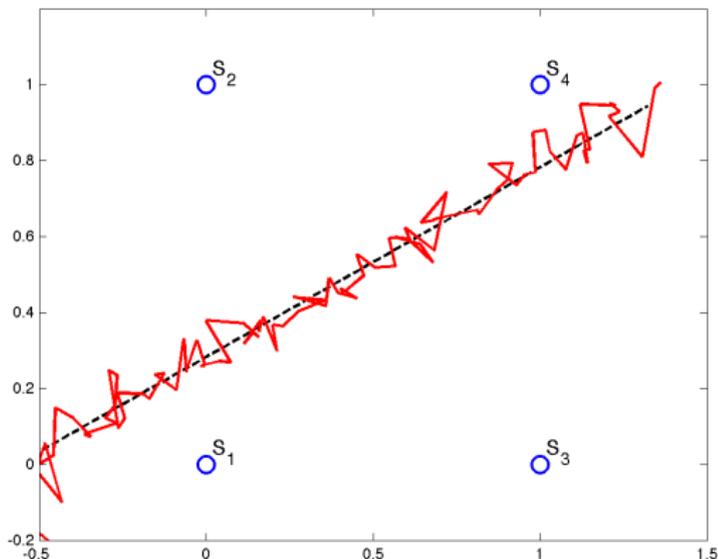
- Resultado de filtrado usando los sensores 1 y 2.



- En este caso los sensores no han podido resolver la ambigüedad de la dirección.

Ejemplo

- Resultado de filtrado usando los cuatro sensores.



Índice

- 1 Un mundo lineal
- 2 Filtro de Kalman extendido
- 3 Filtro de Kalman unscented**
- 4 Un planteamiento más general del problema de estimación
- 5 Monte Carlo
- 6 Importance sampling
- 7 Filtrado de partículas

Filtro de Kalman *unscented*

- Otra extensión no lineal del filtro de Kalman (alternativa al EKF)...
- El modelo:

$$\begin{aligned}\mathbf{x}_t &= \mathbf{f}(\mathbf{x}_{t-1}, \mathbf{v}_t), & \mathbf{v}_t &\sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_n) \\ \mathbf{y}_t &= \mathbf{H}_t \mathbf{x}_t + \mathbf{w}_t, & \mathbf{w}_t &\sim \mathcal{N}(\mathbf{0}, \mathbf{R}_n)\end{aligned}$$

La ecuación de observación es lineal...pero la de estado **no** lo es (\mathbf{f} es una función vectorial arbitraria)

- Se basa en la...

transformación *unscented*

un método para calcular los momentos de una variable aleatoria *gaussiana* tras una transformación no lineal

...que a su vez hace uso de una...

Representación con puntos *sigma*

Sea $p(\mathbf{x}_t | \mathbf{y}_{1:t}) = \mathcal{N}(\mathbf{x}_t | \hat{\mathbf{x}}_t, \mathbf{P}_t)$. Podemos representar esta distribución utilizando un conjunto (determinista) de **puntos sigma**

$$\begin{aligned}\mathbf{X}_t(0) &= \hat{\mathbf{x}}_t, & W_t(0) &= \kappa / (M + \kappa) \\ \mathbf{X}_t(i) &= \hat{\mathbf{x}}_t + \left(\sqrt{(M + \kappa) \mathbf{P}_t} \right)_i, & W_t(i) &= 1 / (2(M + \kappa)) \\ \mathbf{X}_t(i + M) &= \hat{\mathbf{x}}_t - \left(\sqrt{(M + \kappa) \mathbf{P}_t} \right)_i, & W_t(i + M) &= 1 / (2(M + \kappa))\end{aligned}$$

para $i = 1, \dots, M$, donde $\kappa \in \mathbb{R}$ y $\left(\sqrt{(M + \kappa) \mathbf{P}_t} \right)_i$ es la i -ésima columna de la raíz cuadrada (matricial) de $(M + \kappa) \mathbf{P}_t$.

Teorema: Puntos sigma

Este conjunto de muestras ponderadas tiene la misma media muestral y covarianza que la distribución original.

Operación del filtro de Kalman unscented

Una vez calculados los puntos sigma, el **paso de predicción** en el instante $t + 1$ se puede llevar a cabo de la siguiente forma:

- 1 Propagar cada punto sigma mediante la función (no lineal) \mathbf{f}

$$\mathbf{X}_{t+1|t}(i) = \mathbf{f}(\mathbf{X}_t(i), 0).$$

- 2 Calcular la media predictiva

$$\hat{\mathbf{x}}_{t+1}^- = \sum_{i=0}^{2M} W_t(i) \mathbf{X}_{t+1|t}(i).$$

- 3 Calcular la covarianza predictiva

$$\mathbf{P}_{t+1}^- = \sum_{i=0}^{2M} W_t(i) (\mathbf{X}_{t+1|t}(i) - \hat{\mathbf{x}}_{t+1}^-) (\mathbf{X}_{t+1|t}(i) - \hat{\mathbf{x}}_{t+1}^-)^\top$$

El *paso de actualización* es igual que en el FK estándar.

Notas

- El vector de medias y la matriz de covarianza calculadas propagando los puntos sigma a través de la no-linealidad siguen siendo **estimaciones**, pero son más precisas que las dadas por el EKF. Son exactas hasta el orden 2 dado por un desarrollo de Taylor. ✓
- Las aproximaciones todavía son gaussianas, por lo que el método no es adecuado cuando se espera que las distribuciones a posteriori sean multimodales. ✗
- Aunque el UKF no requiere derivadas, conlleva una linealización implícita del modelo (i.e., el UKF se puede reescribir como un método de linealización). ✓
- Los puntos sigma se pueden elegir de otra manera. Por ejemplo, si se utiliza la regla de cuadratura Gauss-Hermite, se necesitan más puntos pero las aproximaciones son más precisas.
- El UKF parece sencillo de implementar. Sin embargo, su rendimiento depende de diversos factores, como por ejemplo el número de puntos.

Índice

- 1 Un mundo lineal
- 2 Filtro de Kalman extendido
- 3 Filtro de Kalman unscented
- 4 Un planteamiento más general del problema de estimación**
- 5 Monte Carlo
- 6 Importance sampling
- 7 Filtrado de partículas

Modelos en espacio de estados

- Planetamiento formal del problema de estimación...
- ...desde el punto de vista bayesiano
- Modelo de espacio de estados no lineal

$$\left\{ \begin{array}{l} \mathbf{x}_t = \mathbf{f}(\mathbf{x}_{t-1}, \mathbf{v}_t) \\ \mathbf{y}_t = \mathbf{h}(\mathbf{x}_t, \mathbf{w}_t) \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} \mathbf{x}_0 \sim p(\mathbf{x}_0) \\ \mathbf{x}_t \sim p(\mathbf{x}_t | \mathbf{x}_{t-1}) \\ \mathbf{y}_t \sim p(\mathbf{y}_t | \mathbf{x}_t) \end{array} \right\}$$

donde

- $\mathbf{f}, \mathbf{h} \equiv$ funciones de transición y observación;
- $\mathbf{v}_t, \mathbf{w}_t \equiv$ ruido del estado y las observaciones;
- $p(\mathbf{x}_0) \equiv$ fdp *a priori* del estado;
- $p(\mathbf{x}_t | \mathbf{x}_{t-1}) \equiv$ fdp de transición del estado;
- $p(\mathbf{y}_t | \mathbf{x}_t) \equiv$ fdp condicional de las observaciones (verosimilitud del estado).

Filtrado estocástico

Objetivo

Hacer seguimiento de la distribución *a posteriori*, $p(\mathbf{x}_t|\mathbf{y}_{1:t})$, que permite calcular la esperanza de cualquier función de interés, \mathbf{g} , como

$$\mathbb{E}[\mathbf{g}(\mathbf{x}_t)] = \int \mathbf{g}(\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t})d\mathbf{x}_t$$

Utilizando el teorema de Bayes, se puede demostrar fácilmente

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) \propto p(\mathbf{y}_t|\mathbf{x}_t) \int p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1})d\mathbf{x}_{t-1}$$

Filtrado estocástico

Hay incertidumbre en las observaciones o en el ruido que determina la evolución del sistema...por eso hablamos de **filtrado estocástico**^a.

^aEl filtro de Kalman también entra en esta categoría!!

Índice

- 1 Un mundo lineal
- 2 Filtro de Kalman extendido
- 3 Filtro de Kalman unscented
- 4 Un planteamiento más general del problema de estimación
- 5 Monte Carlo**
- 6 Importance sampling
- 7 Filtrado de partículas

Integración por Monte Carlo

Siendo X una v.a. con fdp $p(x)$, considere el problema que consiste en aproximar

$$\mathbb{E}[h(X)] = \int h(x)p(x)dx$$

para una función h integrable.



Una manera de abordar el problema

Si somos capaces de **obtener N muestras i.i.d.**, $x^{(1)}, \dots, x^{(N)}$, de $p(x)$ y la varianza de la v.a. $Y = h(X)$ es finita, entonces

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N h(X^{(n)}) = \mathbb{E}[h(X)]$$

casi seguramente (c.s.).

Muestreo

En muchos problemas no es posible directamente obtener muestras de $p(x)$...



Ejemplo

$$\mathbf{y}_t = \mathbf{H}^H \mathbf{x}_t + \mathbf{w}_t$$

Queremos estimar \mathbf{x}_t a partir de \mathbf{y}_t , i.e., queremos aproximar $p(\mathbf{x}_t | \mathbf{y}_t)$...pero no es posible obtener directamente muestras de ella (¿¿cómo??)

...sin embargo, tal vez $p(x)$ puede ser evaluada *hasta una constante de proporcionalidad*²:



$$p(\mathbf{x}_t | \mathbf{y}_t) = \frac{p(\mathbf{y}_t | \mathbf{x}_t)p(\mathbf{x}_t)}{p(\mathbf{y}_t)} \propto p(\mathbf{y}_t | \mathbf{x}_t)p(\mathbf{x}_t)$$

²Digamos $p(x) = Kf(x)$ siendo conocida la función $f(x)$, pero no la constante K .

Índice

- 1 Un mundo lineal
- 2 Filtro de Kalman extendido
- 3 Filtro de Kalman unscented
- 4 Un planteamiento más general del problema de estimación
- 5 Monte Carlo
- 6 Importance sampling**
- 7 Filtrado de partículas

Muestreo enfatizado (Importance sampling)

Supongamos que la fdp de interés, $p(x)$, (la **fdp objetivo**) puede ser evaluada hasta una constante de proporcionalidad y

- *elijamos* una fdp, $q(x)$, conocida como **función tentativa** tal que

$$p(x) > 0 \Rightarrow q(x) > 0$$

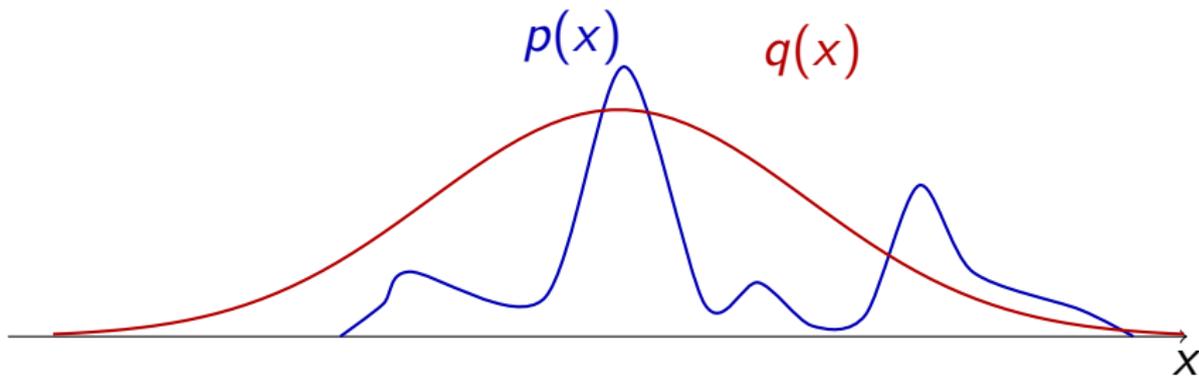
- definamos la función de peso como

$$w(x) = c \frac{p(x)}{q(x)}$$

donde c es una constante arbitraria (posiblemente desconocida)

entonces podemos calcular la esperanza de cualquier función arbitraria $h(x)$ **con respecto a $p(x)$** ...**utilizando muestras de $q(x)$** !!

Importance sampling: $q(x)$



★ Restricción

El soporte de $q(x)$ debe incluir el de $p(x)$,

$$p(x) > 0 \Rightarrow q(x) > 0$$

Cómo elegirla

Por motivos de eficiencia, la función tentativa debería ser la más parecida posible a la objetivo.

Importance sampling: procedimiento

...para aproximar $\mathbb{E}[h(X)]$ respecto de $p(x)$ utilizando muestras de $q(x)$

- 1 Muestrear $\mathbf{x}^{(i)} \sim q(x)$ para $i = 1, \dots, N$
- 2 Calcular

$$w(\mathbf{x}^{(i)}) = c \frac{p(\mathbf{x}^{(i)})}{q(\mathbf{x}^{(i)})} \triangleq w^{*(i)} \text{ (peso sin normalizar)}$$

para $i = 1, \dots, N$

- 3 Normalizar los pesos mediante

$$w^{(i)} = \frac{w^{*(i)}}{\sum_{j=1}^N w^{*(j)}}$$

- 4 Aproximar $\mathbb{E}[h(X)]$ como

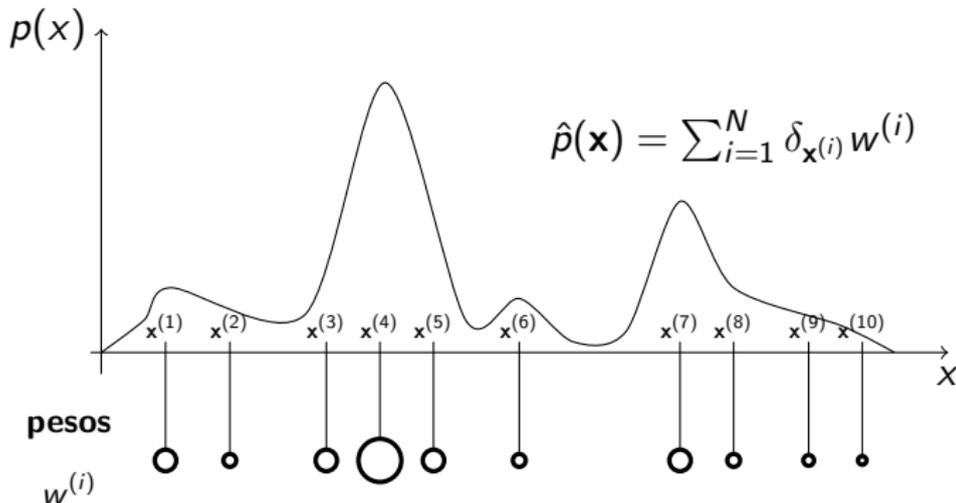
$$\mathbb{E}[h(X)] \approx \sum_{i=1}^N w^{(i)} h(\mathbf{x}^{(i)}) \quad (1)$$

Importance sampling: interpretación

IS da lugar a un conjunto de pares (muestra, peso):

$$\left\{ \left(\mathbf{x}^{(1)}, w^{(1)} \right), \left(\mathbf{x}^{(2)}, w^{(2)} \right), \left(\mathbf{x}^{(3)}, w^{(3)} \right), \dots \right\}$$

El peso se puede *interpretar* como la probabilidad de la muestra correspondiente



Importance sampling en sistemas dinámicos

? IS recursivo

¿Podemos aplicar IS **recursivamente** para estimar el estado en un sistema dinámico?

Consideramos un modelo dinámico en formato de espacio de estados definido por

$$\mathbf{x}_0 \sim p(\mathbf{x}_0), \quad \mathbf{x}_t \sim p(\mathbf{x}_t | \mathbf{x}_{t-1}), \quad \mathbf{y}_t \sim p(\mathbf{y}_t | \mathbf{x}_t)$$

Ya sabemos como aproximar cualquier distribución de interés, por lo que podríamos aproximar

$$p(\mathbf{x}_t | \mathbf{y}_{1:t}), p(\mathbf{x}_{t+1} | \mathbf{y}_{1:t+1}), p(\mathbf{x}_{t+2} | \mathbf{y}_{1:t+2}), \dots$$

pero están relacionadas

Objetivo

Construir (utilizando importance sampling) una aproximación de $p(\mathbf{x}_{t+1} | \mathbf{y}_{1:t+1})$ utilizando una de $p(\mathbf{x}_t | \mathbf{y}_{1:t})$.

Índice

- 1 Un mundo lineal
- 2 Filtro de Kalman extendido
- 3 Filtro de Kalman unscented
- 4 Un planteamiento más general del problema de estimación
- 5 Monte Carlo
- 6 Importance sampling
- 7 Filtrado de partículas**

Importance sampling en sistemas dinámicos

Supongamos que tenemos una aproximación de $p(\mathbf{x}_{t-1} | \mathbf{y}_{1:t-1})$,

$$\hat{p}^N(\mathbf{x}_{t-1} | \mathbf{y}_{1:t-1}) = \sum_{i=1}^N \delta_{\mathbf{x}_{t-1}^{(i)}} w_{t-1}^{(i)}$$

$$\begin{aligned} p(\mathbf{x}_t | \mathbf{y}_{1:t}) &= \frac{p(\mathbf{y}_t | \mathbf{x}_t, \mathbf{y}_{1:t-1})p(\mathbf{x}_t | \mathbf{y}_{1:t-1})}{p(\mathbf{y}_t | \mathbf{y}_{1:t-1})} \\ &\propto p(\mathbf{y}_t | \mathbf{x}_t, \mathbf{y}_{1:t-1})p(\mathbf{x}_t | \mathbf{y}_{1:t-1}) \\ &= p(\mathbf{y}_t | \mathbf{x}_t) \int p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{y}_{1:t-1})p(\mathbf{x}_{t-1} | \mathbf{y}_{1:t-1})d\mathbf{x}_{t-1} \\ &\approx p(\mathbf{y}_t | \mathbf{x}_t) \int p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{y}_{1:t-1})\hat{p}^N(\mathbf{x}_{t-1} | \mathbf{y}_{1:t-1})d\mathbf{x}_{t-1} \\ &= p(\mathbf{y}_t | \mathbf{x}_t) \sum_{i=1}^N p(\mathbf{x}_t | \mathbf{x}_{t-1}^{(i)}, \mathbf{y}_{1:t-1})w_{t-1}^{(i)} \end{aligned}$$

Filtrado de partículas

- Inicialización**

- Se toman muestras de la prior,

$$\mathbf{x}_0^{(i)} \sim p(\mathbf{x}_0), i = 1, \dots, N,$$

- se inicializan todos los pesos al mismo valor

$$w_i^{(0)} = 1/N, i = 1, \dots, N$$

- Recursión** en el instante t

- tomar muestras, $\mathbf{x}_t^{(1)}, \mathbf{x}_t^{(2)}, \dots$, de la función tentativa elegida

$$\mathbf{x}_t^{(i)} \sim q(\mathbf{x}_t | \mathbf{y}_{1:t})$$

- calcular los pesos

$$w_t^{(i)} \propto \frac{p(\mathbf{x}_t^{(i)} | \mathbf{y}_{1:t})}{q(\mathbf{x}_t^{(i)} | \mathbf{y}_{1:t})} = \frac{p(\mathbf{y}_t | \mathbf{x}_t^{(i)}) \sum_{i=1}^N p(\mathbf{x}_t^{(i)} | \mathbf{x}_{t-1}^{(i)}, \mathbf{y}_{1:t-1}) w_{t-1}^{(i)}}{q(\mathbf{x}_t^{(i)} | \mathbf{y}_{1:t})}$$

Este esquema es conocido como **filtrado de partículas** o *importance sampling* secuencial (SIS). Con muestras, se puede utilizar (1) para aprox. cualquier integral respecto de $p(\mathbf{x}_t | \mathbf{y}_{1:t})$.

El filtro *bootstrap*

Si elegimos como función tentativa

$$q(\mathbf{x}_t | \mathbf{y}_{1:t}) = \sum_{i=1}^N p(\mathbf{x}_t | \mathbf{x}_{t-1}^{(i)}, \mathbf{y}_{1:t-1}) w_{t-1}^{(i)}$$

el cálculo de los pesos es muy sencillo

$$\begin{aligned} w_t^{(i)} &\propto \frac{p(\mathbf{x}_t^{(i)} | \mathbf{y}_{1:t})}{q(\mathbf{x}_t^{(i)} | \mathbf{y}_{1:t})} = \frac{p(\mathbf{y}_t | \mathbf{x}_t^{(i)}) \sum_{i=1}^N p(\mathbf{x}_t^{(i)} | \mathbf{x}_{t-1}^{(i)}, \mathbf{y}_{1:t-1}) w_{t-1}^{(i)}}{\sum_{i=1}^N p(\mathbf{x}_t^{(i)} | \mathbf{x}_{t-1}^{(i)}, \mathbf{y}_{1:t-1}) w_{t-1}^{(i)}} \\ &= p(\mathbf{y}_t | \mathbf{x}_t^{(i)}) \end{aligned}$$

El algoritmo resultante es el filtro *bootstrap*, considerado el primer filtro de partículas

El filtro *bootstrap*: la función tentativa

Muestrear de la función tentativa

$$q(\mathbf{x}_t | \mathbf{y}_{1:t}) = \sum_{i=1}^N p(\mathbf{x}_t | \mathbf{x}_{t-1}^{(i)}, \mathbf{y}_{1:t-1}) w_{t-1}^{(i)}$$

se puede interpretar como un procedimiento que consta de dos pasos:

- **remuestrear** la aproximación previa

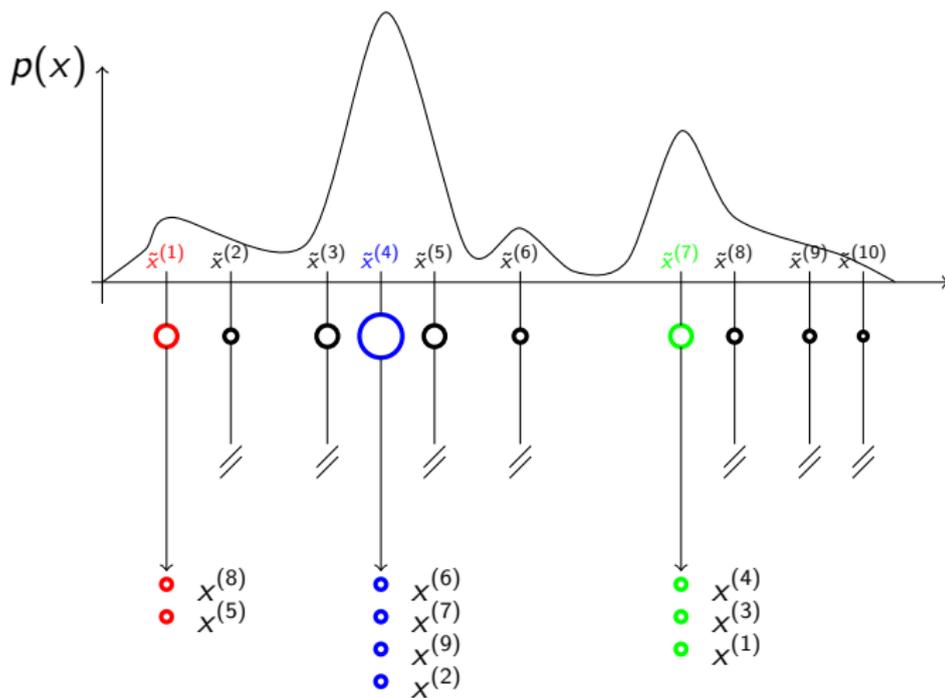
$$\left\{ \left(\mathbf{x}_{t-1}^{(1)}, w_{t-1}^{(1)} \right), \left(\mathbf{x}_{t-1}^{(2)}, w_{t-1}^{(2)} \right), \left(\mathbf{x}_{t-1}^{(3)}, w_{t-1}^{(3)} \right), \dots \right\}$$

para obtener $\mathbf{x}_{t-1}^{(j_1)}, \mathbf{x}_{t-1}^{(j_2)}, \dots, \mathbf{x}_{t-1}^{(j_t)}$ con
 $j_1, j_2, \dots, j_t \in \{1, \dots, N\}$

- **propagar** cada partícula remuestreada utilizando la fdp de transición, $p(\mathbf{x}_t | \mathbf{x}_{t-1})$,

$$\mathbf{x}_t^{(i)} \sim p(\mathbf{x}_t | \mathbf{x}_{t-1}^{(j_i)}), i = 1, \dots, N$$

Remuestreo



El filtro *bootstrap*: implementación

- **Inicialización**

- muestrear $\mathbf{x}_0^{(i)}, i = 1, \dots, N$ de la *prior* $p(\mathbf{x}_0)$

- **Recursión** dado $\hat{p}^N(\mathbf{x}_{t-1} | \mathbf{y}_{1:t-1}) = \sum_{i=1}^N w^{(i)} \delta_{\tilde{\mathbf{x}}_{t-1}^{(i)}}$,

$$\hat{p}^N(\mathbf{x}_{t-1} | \mathbf{y}_{1:t-1}) = \frac{1}{N} \sum_{i=1}^N \delta_{\mathbf{x}_{t-1}^{(i)}}$$

- 1 propagación (muestreo)

$$\tilde{\mathbf{x}}_t^{(i)} \sim p(\mathbf{x}_t | \mathbf{x}_{t-1}^{(i)}), i = 1, \dots, N$$

- 2 cálculo de los pesos...

$$w^{*(i)} = p(\mathbf{y}_t | \tilde{\mathbf{x}}_t^{(i)}), i = 1, \dots, N$$

...y normalización

$$w^{(i)} = \frac{w^{*(i)}}{\sum_{j=1}^N w^{*(j)}}, i = 1, \dots, N$$

- 3 remuestreo: sea $\mathbf{x}_t^{(i)} = \tilde{\mathbf{x}}_t^{(j)}$ con probabilidad $w^{(j)}, i = 1, \dots, N, j \in \{1, \dots, N\}$.

El filtro *bootstrap*: visión general

1. Inicialización

$$\mathbf{x}_0^{(i)} \sim p(\mathbf{x}_0) \text{ para } i = 1, \dots, N$$

2. **Paso recursivo:** partiendo de muestras en el instante $t-1$

2.1. Propagación de muestras

$$\tilde{\mathbf{x}}_t^{(i)} \sim p(\mathbf{x}_t | \mathbf{x}_{t-1}^{(i)})$$

2.2. Cálculo de pesos y normalización

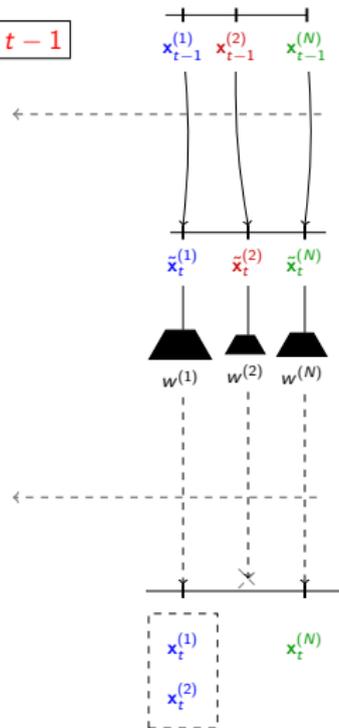
$$w^{(i)} \propto p(\mathbf{y}_t | \tilde{\mathbf{x}}_t^{(i)}), i = 1, \dots, N$$

2.3. Remuestreo

$$\mathbf{x}_t^{(i)} = \tilde{\mathbf{x}}_t^{(j)}, i = 1, \dots, N$$

con probabilidad $w^{(j)}, j \in \{1, \dots, N\}$

muestras en el instante t



El filtro *bootstrap*: epílogo

En la implementación dada arriba, al final de cada iteración tenemos muestras

$$\mathbf{x}_t^{(1)}, \mathbf{x}_t^{(2)}, \dots, \mathbf{x}_t^{(N)}$$

que constituyen una aproximación de

$$p(\mathbf{x}_t | \mathbf{y}_{1:t}),$$

pero el objetivo inicial era aproximar la esperanza de una función de interés (conocida), \mathbf{g} , con respecto a $p(\mathbf{x}_t | \mathbf{y}_{1:t})$, i.e.,

$$\mathbb{E}[\mathbf{g}(\mathbf{x}_t)] = \int \mathbf{g}(\mathbf{x}_t) p(\mathbf{x}_t | \mathbf{y}_{1:t}) d\mathbf{x}_t.$$

Simplemente utilizamos las muestras para obtener una

aproximación por Monte Carlo,

$$\mathbb{E}[\mathbf{g}(\mathbf{x}_t)] \approx \frac{1}{N} \sum_{n=1}^N \mathbf{g}(\mathbf{x}_t^{(n)})$$