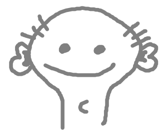
Recursive Estimation

KAHL-MANN: 

Probability review CRV (continuous random variable)

set of possible outcomes: $\mathcal{X}$ (e.g.: $\mathcal{X}=[0,1]$) ; random variable: x ; some value x takes: $\bar{x}\in\mathcal{X}$

probability density func. PDF: $p_x(\bar{x})$ ("density" of probability of event) $(p_x(\bar{x})\geq 0\ \forall\bar{x}\in\mathcal{X}$; $\int_{\mathcal{X}} p_x(\tilde{x})d\tilde{x}=1)$

probability function: $Pr(x\in[a,b])=\int_a^b p_x(\tilde{x})d\tilde{x}$ (probability that $x\in[a,b]$)

cumulative distrib. func. CDF: $F_x(\bar{x})=\int_{-\infty}^{\bar{x}} p(\tilde{x})d\tilde{x}$ (probability that $x<\bar{x}$)

uniform rand. var.: u with $p_u(\bar{u})=\{1$ for $\bar{u}\in(0,1)$, 0 else (e.g. rand in matlab)

---

joint PDF: $p_{xy}(\bar{x},\bar{y})$ (prob. that $x=\bar{x}$ AND $y=\bar{y}$) $(p_{xy}(\bar{x},\bar{y})>0\ \forall\bar{x}\in\mathcal{X},\bar{y}\in\mathcal{Y}$; $\int_{\mathcal{X}}\int_{\mathcal{Y}} p_{xy}(\tilde{x},\tilde{y})\,d\tilde{y}\,d\tilde{x}=1)$

marginalization rule: $p_x(\bar{x})=\int_{\mathcal{Y}} p_{xy}(\bar{x},\tilde{y})d\tilde{y}$ ("sum" over all y prob.) ; $p_{x|z}(\bar{x}|\bar{z})=\int_{\mathcal{Y}} p_{xy|z}(\bar{x},\tilde{y}|\bar{z})d\tilde{y}$

conditioning rule: $p_{x|y}(\bar{x}|\bar{y})=p_{xy}(\bar{x},\bar{y})/p_y(\bar{y})$ (prob($x=\bar{x}$ GIVEN THAT $y=\bar{y}$)) ; $p_{x|yz}(\bar{x}|\bar{y},\bar{z})=\dfrac{p_{xy|z}(\bar{x},\bar{y}|\bar{z})}{p_{y|z}(\bar{y}|\bar{z})}$

usefull relations: $p_x(\bar{x})=\int_{\mathcal{Y}} p_{x|y}(\bar{x}|\tilde{y})\cdot p_y(\tilde{y})d\tilde{y}$; $p_{xy}(\bar{x},\bar{y})=p_y(\bar{y})\cdot p_{x|y}(\bar{x}|\bar{y})=p_x(\bar{x})p_{y|x}(\bar{y}|\bar{x})$

independence: if $p_{x|y}(\bar{x}|\bar{y})=p_x(\bar{x})$ → x,y are independent → $p_{x,y}(\bar{x},\bar{y})=p_x(\bar{x})\cdot p_y(\bar{y})$

conditional independence: if $p_{x|y,z}(\bar{x}|\bar{y},\bar{z})=p_{x|z}(x|z)$ → x,y are cond. indep. on z → $p_{x,y|z}(\bar{x},\bar{y}|\bar{z})=p_{x|z}(\bar{x}|\bar{z})\cdot p_{y|z}(\bar{y}|\bar{z})$

---

expected value: $E_x[x]=\int_{\mathcal{X}}\tilde{x}\cdot p_x(\tilde{x})d\tilde{x}$ (avg. value x is expected to take) $E_{x,y}[f(x,y)]=\int_{\mathcal{Y}}\int_{\mathcal{X}} f(\tilde{x},\tilde{y})\cdot p_{xy}(\tilde{x},\tilde{y})\,d\tilde{x}\,d\tilde{y}$

$E_x[f(x)]=\int_{\mathcal{X}} f(\tilde{x})\cdot p_x(\tilde{x})d\tilde{x}$ (expected value of $f(x)$)

"unconscious statistician": $E_y[y]=\int_{\mathcal{Y}}\tilde{y}p_y(\tilde{y})d\tilde{y}=\int_{\mathcal{X}} g(\tilde{x})p_x(\tilde{x})d\tilde{x}$ where $y=g(x)$

variance: $Var_x[x]=E[(x-E[x])^2]=E[x^2]-E[x]^2$ (squared avg. dist. from $E[x]$) $Var_{\underline{x}}[\underline{x}]=E[(\underline{x}-E[\underline{x}])\cdot(\underline{x}-E[\underline{x}])^T]$; $Var_x[f(x)]=E_x[f(x)^2]-E_x[f(x)]^2$

affine transform: $y=mx+b$ → $E_y[y]=mE_x[x]+b$; $Var_y[y]=m^2\cdot Var_x[x]$

$\underline{y}=\underline{\underline{M}}\underline{x}+\underline{b}$ → $E_{\underline{y}}[\underline{y}]=\underline{\underline{M}}E_{\underline{x}}[\underline{x}]+\underline{b}$; $Var_{\underline{y}}[\underline{y}]=\underline{\underline{M}}\,Var_{\underline{x}}[\underline{x}]\,\underline{\underline{M}}^T$

---

sample PDF: $p_x(\bar{x})\to F_x(\bar{x})\to F_x(x)=u\to\bar{x}=F_x^{-1}(\bar{u})$ (u: uniform rand. var. from e.g. matlab)

sample joint PDF: $p_{xy}(\bar{x},\bar{y})\to p_x(\bar{x})\to F_x(\bar{x})\to F_x(x)=u_1\to\bar{x}=F_x^{-1}(\bar{u}_1)$
$\searrow p_{y|x}(\bar{y}|\bar{x})\to F_{y|\bar{x}}(\bar{y}|\bar{x})\to F_{y|\bar{x}}(y|\bar{x})=u_2\to\bar{y}=F_{y|\bar{x}}^{-1}(\bar{u}_2|\bar{x})$ ($\bar{u}_1,\bar{u}_2$: 2 uniform independ. vars.)

change of variable: $\bar{y}=g(\bar{x})$: $p_y(\bar{y})=p_x(g^{inv}(\bar{y}))/|\frac{dg}{dx}(g^{inv}(\bar{y}))|$

$(\bar{y}_1,\bar{y}_2,..)=G(\bar{x}_1,\bar{x}_2,..)$: $p_{\underline{y}}(\bar{\underline{y}})=p_{\underline{x}}(G^{inv}(\bar{\underline{y}}))/|J_G(G^{inv}(\bar{\underline{y}}))|$ with $|J_G|=\left|\det\begin{pmatrix}\partial G_1/\partial x_1 & \cdots & \partial G_n/\partial x_1\\ \vdots & & \vdots\\ \partial G_1/\partial x_n & \cdots & \partial G_n/\partial x_n\end{pmatrix}\right|$

$\bar{\underline{y}}=\underline{\underline{M}}\cdot\bar{\underline{x}}+\underline{b}$: $p_{\underline{y}}(\bar{\underline{y}})=p_{\underline{x}}(\underline{\underline{M}}^{-1}(\bar{\underline{y}}-\underline{b}))/|\det(\underline{\underline{M}})|$

---

gaussian rand. val.: $p(\underline{y})=1/2\pi^{D/2}\cdot\det(\underline{\underline{\Sigma}})^{1/2}\cdot\exp\left(-\frac{1}{2}(\underline{y}-\underline{\mu})^T\underline{\underline{\Sigma}}^{-1}(\underline{y}-\underline{\mu})\right)$ $\underline{\mu}\in\mathbb{R}^D$: mean vector $\underline{\underline{\Sigma}}\in\mathbb{R}^{D\times D}$: covariance matrix

if mutually indep. val.: $p(\underline{y})=\prod_{i=1}^{D} 1/\sqrt{2\pi\sigma_i^2}\cdot\exp\left(-\frac{1}{2\sigma_i^2}\cdot(y_i-\mu_i)^2\right)$ $\underline{\underline{\Sigma}}=\begin{bmatrix}\sigma_1^2 & & 0\\ & \ddots & \\ 0 & & \sigma_D^2\end{bmatrix}$; σ_i^2: variances

jointly gaussian RV: if $p(\underline{x})$ and $p(\underline{y})$ gaussian + independent, then $p(\underline{x},\underline{y})$ also gaussian: $p(x,y)\propto\exp\left(-\frac{1}{2}\begin{bmatrix}x-\mu_x\\ y-\mu_y\end{bmatrix}^T\begin{bmatrix}\Sigma_x^{-1} & 0\\ 0 & \Sigma_y^{-1}\end{bmatrix}\begin{bmatrix}x-\mu_x\\ y-\mu_y\end{bmatrix}\right)$

affine transform: $\underline{y}=\underline{\underline{M}}\underline{x}+\underline{b}$ → $\mu_y=\underline{\underline{M}}\mu_x+\underline{b}$; $\underline{\underline{\Sigma}}_y=\underline{\underline{M}}\underline{\underline{\Sigma}}_x\underline{\underline{M}}^T$ (affine transforms preserve GRV form)

---

particles from PDF: **sample** PDF $p_x(\bar{x})$ N **times** producing particles $\bar{x}_n$ $(n=1,\dots,N)$ (density of particles ≙ PDF)

PDF from particles: $p_x(\bar{x})\approx\frac{1}{N}\sum_{n=1}^{N}\delta(\bar{x}-\bar{x}_n)$ (δ: Dirac delta for CRV Kronecker delta for DRV)

transforms with particles: given $\bar{y}=g(\bar{x})$ → $\bar{y}_n=g(\bar{x}_n)$ for $n=1,\dots,N$ (transform every particle to get new particles)

resample particles: given $p_y(\bar{x}_n)$ → $\bar{y}_n=$ **sample** DRV $p_y(\bar{x}_n)$ N **times** ($\bar{y}_n$ will be a subset of $\bar{x}_n$ with PDF p_y)

Probability review DRV

(discrete random variable) same as CRV except:

1) $\mathcal{X}$ finite set of values   2) $p_x(\bar{x})$: probability of picking value $\bar{x}$ 3) integrals over $\mathcal{X}$ become summs over $\mathcal{X}$

Bayes' Theorem

$$p(x|z) = \frac{p(z|x)p(x)}{p(z)}$$

x: unknown system state
z: observation of state

$p(x)$: prior belief of state (before observing, what is prob. of x ?)
$p(z|x)$: observation model (for given x, what is prob. of observing z)
$p(x|z)$: posteriori belief of state (for given observation z, what is prob. of state x)
$p(z) = \int_\mathcal{X} p(z|\tilde{x})\, p(\tilde{x})\, d\tilde{x}$: normalization constant

$$p(x|z_1..z_N) = \frac{p(x)\prod_{i=1}^{N} p(z_i|x)}{\int_\mathcal{X} p(\tilde{x})\prod_{i=1}^{N} p(z_i|\tilde{x})\,d\tilde{x}}$$ (form for multiple z_i observations that are conditionally independent on x)

Bayesian Tracking

state model: $p(x(k)|x(k-1))$ posteriori state belief: $p(x(k)|z(1{:}k))$ (PDF of x(k) given obs. at times 1,...,k)
observ. model: $p(z(k)|x(k))$ priori state belief: $p(x(k)|z(1{:}k-1))$ (PDF of x(k) given obs. at times 1,...,k-1)

- **priori update:**

leave out since x(k) and z(1:k-1) are condit. indep. on x(k-1) (lecture 4 for proof...)

$$\underbrace{p(x(k)|z(1{:}k-1))}_{\text{prior}} = \int_\mathcal{X} \underbrace{p(x(k)|x(k-1),\cancel{z(1{:}k-1)})}_{\text{state model}}\, \underbrace{p(x(k-1)|z(1{:}k-1))}_{\text{prev. posterior}}\, dx(k-1)$$

(• known, • integ. var., • result var.)

- **posteriori update:**

new measurement; leave out since z(k) and z(1:k-1) are condit. indep. on x(k)

$$\underbrace{p(x(k)|z(1{:}k))}_{\text{posterior}} = p(x(k)|z(k),z(1{:}k-1)) = \underbrace{p(z(k)|x(k),\cancel{z(1{:}k-1)})}_{\text{observ. model}} \cdot \underbrace{p(x(k)|z(1{:}k-1))}_{\text{prior}} \,/\, \underbrace{p(z(k)|z(1{:}k-1))}_{\text{normalization}}$$

$$p(z(k)|z(1{:}k-1)) = \int_\mathcal{X} \underbrace{p(z(k)|x(k),\cancel{z(1{:}k-1)})}_{\text{observ. model}} \cdot \underbrace{p(x(k)|z(k-1))}_{\text{prior}}\, dx(k)$$

- **discrete state space implementation:**

$a^i_{0|0} = p_{x(0)}(i)$ $(i=0,..,N-1)$

$a^i_{k|k-1} = \sum_{j=0}^{N-1} p_{x(k)|x(k-1)}(i|j)\, a^j_{k-1|k-1}$ $(i=0,..,N-1)$

$$a^i_{k|k} = \frac{p_{z(k)|x(k)}(\bar{z}(k)|i)\, a^i_{k|k-1}}{\sum_{j=0}^{N-1} p_{z(k)|x(k)}(\bar{z}(k)|j)\, a^j_{k|k-1}}$$ $(i=0,..,N-1)$

$\mathcal{X} = \{0,..,N-1\}$
$a^i_{k|k} = p_{x(k)|z(1:k)}(i|\bar{z}(1{:}k))$ posteriori discrete PDF $(i=0,..,N-1)$
$a^i_{k|k-1} = p_{x(k)|z(1:k-1)}(i|\bar{z}(1{:}k-1))$ priori discrete PDF $(i=0,..,N-1)$
$p_{x(k)|x(k-1)}(i|j)$: state model ; $p_{z(k)|x(k)}(\bar{z}(k)|j)$: observ. model

- **observation model:**

(x=x(k) : system state
z=z(k) : observation
w=w(k) : obs. noises)

$z = h(x,w)$
$w = h^{inv}(x,z)$
$|J_h|(x,w) = |\det(\partial h_i/\partial w_j)|$

$p_{z|x}(\bar{z}|\bar{x})$ = prob. of observing $\bar{z}$, given state $\bar{x}$
$= p_{w|x}(\bar{w}|\bar{x}) \,/\, |J_h|(\bar{x},\bar{w})$ (change of variable)
$= p_w(\bar{w}) \,/\, |J_h|(\bar{x},\bar{w})$ (w,x are independ.)
$= p_w(h^{inv}(\bar{x},\bar{z})) \,/\, |J_h|(\bar{x},h^{inv}(\bar{x},\bar{z}))$ (h is invertible in w)

- **state model:**

(x(k): system state
v(k): sys. noises)

$x(k) = g(x(k-1),v(k-1))$
(analog to obs. model)

$p(x(k)|x(k-1))$ = prob. next state $\bar{x}(k)$, given x(k-1)
$= p_v(g^{inv}(x(k-1),x(k))) \,/\, |J_g|(x(k-1),g^{inv}(x(k-1),x(k)))$

- **estimates of state:**

– maximum likelihood (ML): $\hat{x}^{ML} = \underset{\bar{x}\in\mathcal{X}}{\text{argmax}}\ p_{z|x}(\bar{z}|\bar{x})$ ($p_{z|x}(\bar{z}|\bar{x})$: observation model)
– maximum a posteriori (MAP): $\hat{x}^{MAP} = \underset{\bar{x}\in\mathcal{X}}{\text{argmax}}\ p_{x|z}(\bar{x}|\bar{z}) = \underset{\bar{x}\in\mathcal{X}}{\text{argmax}}\ p_{z|x}(\bar{z}|\bar{x})\cdot p_x(\bar{x})$ ($p_{x|z}(\bar{x}|\bar{z})$: a posteriori belief)
– recursive least square (RLS): same as only posteriori update of Kalmann-Filter

Kalman-Filter (Bayesian tracking with: linear models, gaussian noise → gausian state beliefs)

state model: $p(x(k)|x(k\text{-}1))$ priori state belief: $p_{x_p}(k)=p(x(k)|z(1{:}k\text{-}1))$ (x_p rand.var.; $\hat{x}_p$ mean; P_p variance)

observ. model: $p(z(k)|x(k))$ posteriori state belief: $p_{x_m}(k)=p(x(k)|z(1{:}k))$ (x_m rand.var.; $\hat{x}_m$ mean; P_m variance)

system models: $x(k)=A\cdot x(k\text{-}1)+Bu(k\text{-}1)+v(k\text{-}1)$

$z(k)=H\cdot x(k)+w(k)$

gaussian noises: $p_{x_m}(0)$ ($x_m(0)$ rand.var.; x_0 mean; P_0 variance)

$p(v(k))$ (v rand.var.; 0 mean; Q variance)

$p(w(k))$ (w rand.var.; 0 mean; R variance)

- **priori update**: $p_{x_p}(k)=\int_{\chi} p(x(k)|x(k\text{-}1))\cdot p_{x_m}(k\text{-}1)\,dx(k\text{-}1)$

  hard to work with! ↓ lecture 6 for proof...

  better approach: use formula $x_p(k)=Ax_m(k\text{-}1)+B\,u(k\text{-}1)+v(k\text{-}1)$

  → $\hat{x}_p(k)=E[x_p(k)]=A\cdot E[x_m(k\text{-}1)]+u(k\text{-}1)+E[v(k\text{-}1)]=A\hat{x}_m(k\text{-}1)+Bu(k\text{-}1)$

  → $P_p(k)=\text{Var}[x_p(k)]=E[(x_p(k)-\hat{x}_p(k))\cdot(x_p(k)-\hat{x}_p(k))]=\ldots=AP_m(k\text{-}1)A^T+Q$

- **posteriori update**: $p_{x_m}(\bar{x}_m)\propto \underbrace{p(\bar{z}|\bar{x}_m)}\cdot\underbrace{p_{x_p}(\bar{x}_m)}$ ← k skipped for shortness

  $p_{x_m}(\bar{x}_m)\propto\exp(-\frac{1}{2}((\bar{z}-H\bar{x}_m)^T R^{-1}(\bar{z}-H\bar{x}_m)+(\bar{x}_m-\hat{x}_p)^T P_p^{-1}(\bar{x}_m-\hat{x}_p)))=\exp(-\frac{1}{2}(\bar{x}_m-\hat{x}_m)^T P_m^{-1}(\bar{x}_m-\hat{x}_m))$

  → $P_m(k)=\ldots=(P_p(k)^{-1}+H^TR^{-1}H)^{-1}$

  → $\hat{x}_m(k)=\ldots=\hat{x}_p(k)+P_m(k)H^TR^{-1}(\bar{z}(k)-H\hat{x}_p(k))$

- $p(x(k)|x(k\text{-}1))=$ state model

  $\propto\exp(-\frac{1}{2}(x(k)-Ax(k\text{-}1)-Bu(k\text{-}1))^TQ^{-1}(x(k)-Ax(k\text{-}1)-Bu(k\text{-}1)))$

  $x(k)=g(\ldots)=Ax(k\text{-}1)+B\,u(k\text{-}1)+v(k\text{-}1)$

  $v(k\text{-}1)=g^{inv}(\ldots)=x(k)-Ax(k\text{-}1)-Bu(k\text{-}1)$

  $|J_g|=|\det(\partial g_i/\partial v_j)|=1$

  $p_v(\bar{v})\propto\exp(-\frac{1}{2}\cdot\bar{v}^TQ^{-1}\bar{v})$

- $p(z(k)|x(k))=$ observ. model

  $\propto\exp(-\frac{1}{2}(z(k)-Hx(k))^TR^{-1}(z(k)-Hx(k)))$

  ↑ analog to state model

- all equations for **posteriori update**:

  $K(k)=P_p(k)H^T\cdot(HP_p(k)H^T+R)^{-1}$ (Kalman Filter gain)

  $\hat{x}_m(k)=\hat{x}_p(k)+K(k)(\bar{z}(k)-H\hat{x}_p(k))$

  $P_m(k)=(I-K(k)H)P_p(k)$ [$=(I-K(k)H)P_p(k)(I-K(k)H)^T+K(k)RK(k)^T$: Joseph form, more comp. expensive, but less numerical error]

  (A,H,Q,R can be func of k. if they are const./known then $K(k), P_m(k), P_p(k)$ can be computed offline!)

Steady state Kalmann-Filter ($P_p(k), P_m(k), K(k)$ might converge for $k\to\infty$. can simplify implementation)

$P_\infty=\lim_{k\to\infty}P_p(k)$ → $P_\infty=AP_\infty A^T+Q-AP_\infty H^T(HP_\infty H^T+R)^{-1}HP_\infty A^T$

($P_p(k)=P_p(k+1)$ → Riccati eq. → solve for $P_\infty=\ldots$)

$K_\infty=\lim_{k\to\infty}K(k)$ → $K_\infty=P_\infty H^T(HP_\infty H^T+R)^{-1}$

→ $\hat{x}_m(k)=\underbrace{(I-K_\infty H)A}_{\hat{A}}\hat{x}_m(k\text{-}1)+\underbrace{(I-K_\infty H)B}_{\hat{B}}u(k\text{-}1)+K_\infty\bar{z}(k)$

$\hat{x}_m(k)=\hat{A}\cdot\hat{x}_m(k+1)+\hat{B}\cdot u(k\text{-}1)+K_\infty\bar{z}(k)$

[alternative formulation: $\hat{x}_m(k)=A\hat{x}_m(k\text{-}1)+Bu(k\text{-}1)+K_\infty(\bar{z}(k)-\hat{z}(k))$; $\hat{z}(k)=H(A\hat{x}(k\text{-}1)+Bu(k\text{-}1))$ (Luenberger form)]

[error dynamics: $e(k)=x(k)-x_m(k)=\ldots=(I-K_\infty H)Ae(k\text{-}1)+(I-K_\infty H)v(k\text{-}1)-K_\infty w(k)$ → $E[e(k)]=(I-K_\infty H)A\cdot E[e(k\text{-}1)]$]

necessary presumptions:

- P_∞ must exist and be positive semidefinite and constant for any initial condition $P_p(1)$

  e.g. SISO rank 1 system: $P_\infty=\frac{a^2rP_\infty}{h^2P_\infty+r}+q$

  - $|a|<1$: one solution for any $P_p(1)$
  - $|a|\geq1, h\neq0, q>0$: one solution for any $P_p(1)$
  - $|a|\geq1, h=0, q>0$: no solution
  - $|a|\geq1, h=0, q=0$: one solution $P_\infty=0$ for $P_p(1)=0$ (unstable)
  - $|a|\geq1, h\neq0, q=0$: two solutions, one like ↑, one for other $P_p(1)$
- error must go to 0 ↔ error dynamics must be stable ↔ $|\lambda_i|<1\ \forall i$; $\lambda_i=EW((I-K_\infty H)A)$

alternative (equivalent) presumptions:

- (A,H) is detectable: all unobservable states are stable $\quad \text{rank}(A^T-\lambda_iI \,\vdots\, H^T)=n \quad \forall|\lambda_i|\geq1;\ \lambda_i=EW(A)$
- (A,G) is pot. stabilizable: all noiseless states are stable $\quad \text{rank}(\lambda_iI-A \,\vdots\, G)=n \quad \forall|\lambda_i|\geq1;\ \lambda_i=EW(A)$

  ↖ $Q=GG^T$

Extended Kalman-Filter (extension for nonlinear models. linearize at each timestep)

• priori update: (input u is integrated in $q(x,v)$)

$x(k)=q(x(k-1),v(k-1))$ linearize at $x=\hat{x}_m(k-1)$; $v=E[v(k-1)]=0$:

$$\approx q(\hat{x}_m(k-1),0)+\overbrace{\frac{\partial q}{\partial x}(\hat{x}_m(k-1),0)}^{A(k-1)}\cdot(x(k-1)-\hat{x}_m(k-1))+\overbrace{\frac{\partial q}{\partial v}(\hat{x}_m(k-1),0)}^{L(k-1)}\cdot v(k-1)$$

$$=A(k-1)x(k-1)+\underline{L(k-1)v(k-1)}+\underline{q(\hat{x}_m(k-1),0)-A(k-1)\hat{x}_m(k-1)}$$

$$=A(k-1)x(k-1)+\tilde{v}(k-1)+\xi(k-1)$$

↓ analog derivation to ordinary Kalman-Filter:

$$\hat{x}_p(k)=A(k-1)\hat{x}_m(k-1)+\xi(k-1)=\ldots=q(\hat{x}_m(k-1),0)$$

$$P_p(k)=A(k-1)P_m(k-1)A^T(k-1)+L(k-1)Q(k-1)L^T(k-1)$$

$\boxed{\text{EKF only approximates } \hat{x}_p,\hat{x}_m,P_p,P_m \rightarrow \text{no guarantees!}}$

• posteriori update:

$z(k)=h(x(k),w(k))$ linearize at $x=\hat{x}_p(k)$; $w=E[w(k)]=0$:

$$\approx h(\hat{x}_p(k),0)+\overbrace{\frac{\partial h}{\partial x}(\hat{x}_p(k),0)}^{H(k)}\cdot(x(k)-\hat{x}_p(k))+\overbrace{\frac{\partial h}{\partial w}(\hat{x}_p(k),0)}^{M(k)}\cdot w(k)$$

$$=H(k)x(k)+\underline{M(k)w(k)}+\underline{h(\hat{x}_p(k),0)-H(k)\hat{x}_p(k)}$$

$$=H(k)x(k)+\tilde{w}(k)+\zeta(k)$$

↓ analog derivation to ordinary Kalman-Filter:

$$K(k)=P_p(k)H^T(k)\left(H(k)P_p(k)H^T(k)+M(k)R(k)M^T(k)\right)^{-1}$$

$$\hat{x}_m(k)=\hat{x}_p(k)+K(k)(\bar{z}(k)-H(k)\hat{x}_p(k)-\zeta(k))$$

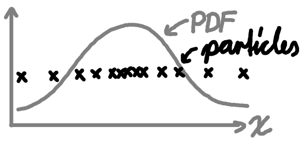
$$=\hat{x}_p(k)+K(k)(\bar{z}(k)-h(\hat{x}_p(k),0))$$

$$P_m(k)=(I-K(k)H(k))P_p(k)$$

Hybrid Kalman-Filter (like EKF, but with continuous time state model) (don't use this. just discretize model!)

Particle Filter (Bayesian tracking with particles as PDFs)

use particles to approximate PDFs
where particle density ≈ prob. density
(particles generated by sampling PDF)

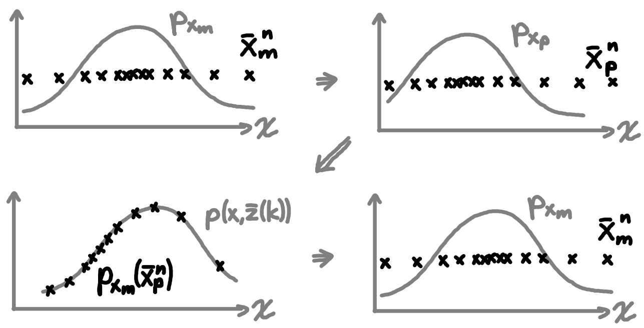


$\bar{x}_m^n(k)$: $n=1,\ldots,N$ particles = samples of $p_{x_m}(k)$

$\bar{x}_p^n(k)$: $n=1,\ldots,N$ particles = samples of $p_{x_p}(k)$

• priori update: use $x_p(k)=q(x_m(k-1),u(k-1),v(k-1))$ on every particle of x_m to get new particles for x_p

$$\bar{x}_p^n(k)=q(\bar{x}_m^n(k-1),u(k-1),\bar{v}^n(k-1)) \quad (\bar{v}^n(k-1)\text{: noise samples})$$



• posteriori update:

$$p_{x_m}(\bar{x}_m)\propto p(\bar{z}|\bar{x}_m)\cdot p_{x_p}(\bar{x}_m)=p(\bar{z}|\bar{x}_m)\cdot\sum_{n=1}^{N}\delta(\bar{x}_m-\bar{x}_p^n)$$

$\rightarrow p_{x_m}(\bar{x}_p^n)\propto p(\bar{z}|\bar{x}_p^n)=p(\bar{z}|\bar{x}_p^n)/\left(\sum_{i=1}^{N}p(\bar{z}|\bar{x}_p^i)\right)$ (normalize)

$\rightarrow \bar{x}_m^n=$ sample DRV $p_{x_m}(\bar{x}_p^n)$ N times $+g^n$ (optional)

{ might want to add noise here as $\bar{x}_m^n$ is a subset of $\bar{x}_p^n$, which can lead to sample impoverishment (all samples in one place) → particles no longer approx PDF

e.g.: $\mu_g=0$, $\sigma_g^i=\sqrt{var_g}=KE_iN^{-1/d}$ { K: tuning param; E_i = max particle spread; N: # of particles; d: # dimensions of x

Separation Principle

for LTI systems with gaussian noise the steady-state Kalman-filter is the optimal state observer! and an LQR controller $u(k)=F_\infty x(k)$ is optimal in reducing a quadratic cost! combining the two yields Linear-Quadratic-Gaussian (LQG) control which is the globally optimal control strategy (= separation theorem).

state model: $x(k)=Ax(k-1)+Bu(k-1)+v(k-1)$

observer model: $z(k)=Hx(k)+w(k)$

KF observer: $\hat{x}(k)=(I-K_\infty H)A\hat{x}(k-1)+(I-K_\infty H)Bu(k-1)+K_\infty\bar{z}(k)$

LQR controller: $u(k)=F_\infty\hat{x}(k)$

error dynamics: $e(k)=(I-K_\infty H)A\cdot e(k-1)+(I-K_\infty H)v(k-1)-K_\infty w(k)$ $(e(k)=x(k)-\hat{x}(k))$

state dynamics: $x(k)=Ax(k-1)+BF_\infty\hat{x}(k-1)+v(k-1)=(A+BF_\infty)x(k-1)+BF_\infty e(k-1)+v(k-1)$

↳ combined system (for means):

$$\begin{bmatrix}E[x(k)]\\E[e(k)]\end{bmatrix}=\begin{bmatrix}A+BF_\infty & -BF_\infty\\0 & (I-K_\infty H)A\end{bmatrix}\cdot\begin{bmatrix}E[x(k-1)]\\E[e(k-1)]\end{bmatrix}$$

→ { for system to be stable $(I-K_\infty H)A$ and $(A+BF_\infty)$ must be stable. This is equivalent to (A,H) being detectable + (A,B) pot. stabilizable