EX: $P(X_n = s_n \mid X_{0:n-1} = s_{0:n-1}) = P(X_n = s_n \mid X_{n-1} = s_{n-1})$ [Def. of a M.C]

$X_n = f_n(X_{n-1}, U_n)$; $\{U_n\}$ indep. } of: $X_0, \overline{U_1, \ldots, U_{n-1}}$

$$\begin{bmatrix} X_1 = f_1(X_0, U_1) \\ \vdots \\ X_{n-1} = f_{n-1}(X_{n-2}, U_{n-1}) \end{bmatrix}$$

$\Downarrow$

$\{X_n\}$ is Markov

$$P[X_n = s_n \mid X_{0:n-1} = s_{0:n-1}] = \sum_{u:\, s_n = f_n(s_{n-1}, u)} P[U_n = u \mid X_{0:n-1} = s_{0:n-1}]$$

$P\{\omega : X_n(\omega) = s_n\}$

$P(X/A) = P(X)$

$$= \sum_{u:\, s_n = f_n(s_{n-1}, u)} P[U_n = u] = f \cdot \text{ of } s_{n-1}$$

only $\to$ M.C

$s_k \in \mathcal{S} = \{\omega_1, \ldots, \omega_r\}$

1/ * Algorithmic Properties of M.C. $\to$ Compute efficiently filters (Estimators)

2/ * Asymptotic properties (Ergodic) $\to$ Simulation & optimization

1/ Dynamic programming. Signal processing & process. controlling

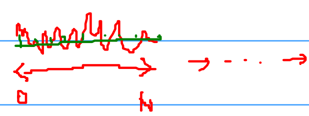
$$\begin{cases} X_n : \text{Markov} \\ W_n : \text{Noise} \\ Y_n = h_n(X_n, W_n) \end{cases}$$

$\{W_n\}$ independent Seq. (White Noise)

$n = 0 : N$

X_0

X_{-1}



$$\underbrace{J_N(s_{0:N})} = P(Y_{0:N} ; \overline{s_{0:N}}) = \underbrace{P(Y_{0:N} / X_{0:N} = s_{0:N})}\ \underbrace{P(s_{0:N})}$$

$$\prod_{k=0}^{N} P(Y_k \mid \underbrace{X_k = s_k}) * P_0(s_0)\, P(s_2 \mid s_1) \cdots P(s_n \mid s_{n-1}) \cdots P(s_N \mid s_{N-1})$$

$B_k(s_k)$

$P(s_1 \mid s_0)$

$P(s_N \mid s_{0:N-1})\, \underbrace{P(s_{0:N-1})}$

$\Downarrow$ $P(s_N \mid s_{N-1})$; $P(s_{N-1} \mid s_{0:N-2})\, P(s_{0:N-2})$

$\Downarrow$ $P(s_{N-1} / s_{N-2}) \cdots$

$$= P_X(s_0) \cdots g_n(s_{n-1}, s_n) \cdots g_N(s_{N-1}, s_N)$$

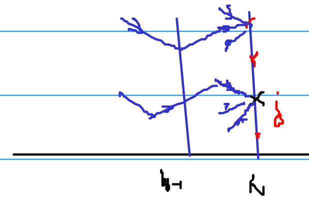
$g_{n+1}(s_n, s_{n+1})$

Problem: find $\left\{ \hat{s}_{0:N} = \underset{s_{0:N}}{\operatorname{ArgMax}}\ J_N(s_{0:N}) \right.$

Dynamic Programming

* Let $\hat{J}_N = \max_{S_{0:N}} J_N(S_{0:N})$ $\qquad S_k \in \mathcal{S} = \{1,\dots,r\}$

$$= \max_j \bar{J}_N(j) = \max_j \max_{S_{0:N-1}} J_N(S_{0:N-1}, S_N = j)$$

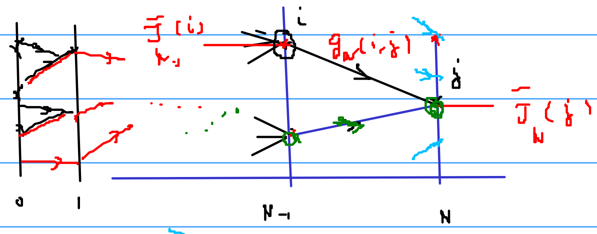


Recursion for $\bar{J}_N(j)$

$$\bar{J}_N(j) = \max_i \max_{S_{0:N-2}} J_N(S_{0:N-2}, S_{N-1} = i, S_N = j) \leftarrow g_0(x_0)\cdots g_{N-1}(S_{N-2}, S_{N-1})\, g_N(S_{N-1}, S_N)$$

$$= \max_i \underbrace{g_N(i,j)\, \bar{J}_{N-1}(i)}_{L_N(i,j)}$$

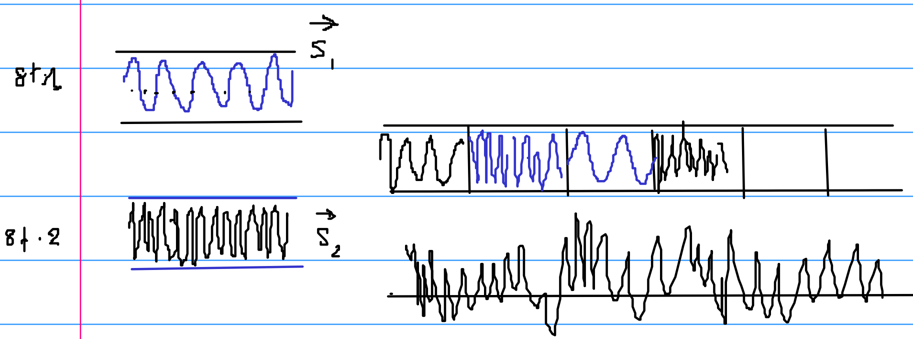
$$\hat{S}_{N-1}(j) = \operatorname{Arg\,max}_i L_N(i,j)$$



$|\mathcal{S}|^N \to N$

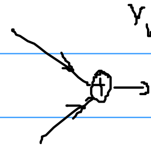
$g_n(i,j) = P(Y_n \mid S_n = j)$

$p(S_n = j \mid S_{n-1} = i)$



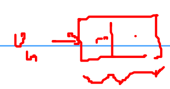
$\vec{Y}_n = f_n(\vec{X}_n, \vec{W}_n)$

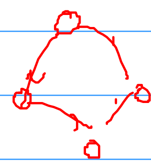
$\vec{Y}_n = h_n(\vec{X}_{n-1}, \vec{X}_n, \vec{W}_n)$



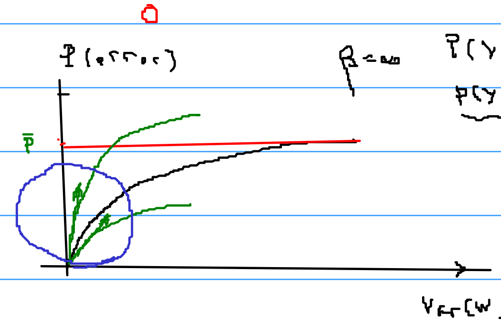
Y_n $\qquad Z_n = \begin{bmatrix} X_{n-1} \\ X_n \end{bmatrix}$

$\bar{Y}_n = h_n'(Z_n, Y_n)$





Particle Filters : $\left\{\begin{array}{l}\text{In finite} \\ \text{no. of states.}\end{array}\right.$



$\beta = \infty$ $\qquad P(Y, X)$

$p(Y|X)\, P(x)$

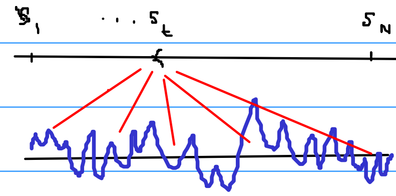
$(x_{1:N}) \to (y_{1:N}) \xrightarrow{?} \max_{\underline{s}} P(\underline{y}, \underline{s}) \to$ Min. Block Error $\to$ Viterbi (Dyn. Prog.)

$$\left\{ \max_{s_t} \underbrace{P(s_t \mid y_{1:N})} \right. \quad t = 1:N$$

$\to$ Min Prob. Symbol error

Speech processing.

* Forward-Backward Algorithm (BCJR)



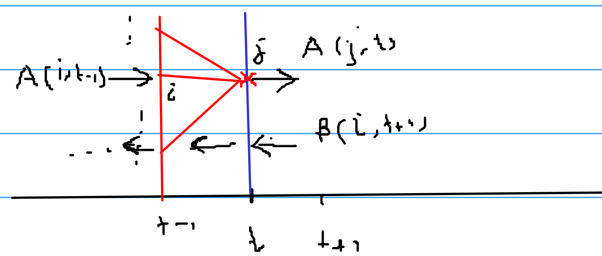
$$P(s_t = j \mid y_{1:N}) = P(s_t = j \,;\, y_{1:N}) / P(y_{1:N})$$

$$\propto \sum_{s_{1:t-1},\, t+1:N} P(s_{1:t-1},\, s_t = j,\, s_{t+1:N} \,;\, y_{1:N})$$

$$\sum_{s_{1:t-1},\, s_{t+1:N}} P(s_1) \cdots P(y_t \mid s_t)\, P(\underbrace{s_t}_{s_t = j} \mid s_{t-1}) \cdots P(y_N \mid s_N)\, P(s_N \mid s_{N-1})$$

$$\underbrace{\sum_{s_{1:t-1}} P(s_1) \cdots P(y_t \mid s_t)\, P(s_t = j \mid s_{t-1})}_{A(j,t)} \; \underbrace{\sum_{s_{t+1:N}} P(s_{t+1} \mid s_t = j)\, P(y_{t+1} \mid s_{t+1}) \cdots P(y_N \mid s_N)\, P(s_N \mid s_{N-1})}_{B(j,t)}$$

$$\begin{cases} A(j,t) = \sum_i A(i, t-1)\, P(y_t \mid s_t = j)\, P(s_t = j \mid s_{t-1} = i) \\ B(j,t) = \sum_i P(s_{t+1} = j \mid s_t = i)\, P(y_{t+1} \mid s_{t+1} = j)\, B(i, t+1) \end{cases}$$



Gaussian Markov process :

Evol. Eq. $\rightarrow$ $M \left\{ \vec{X}_n = A_n \vec{X}_{n-1} + \vec{U}_n \right.$, $\underbrace{X_0, U_0, \ldots, U_n}_{\text{Independent}}$ Jointly gaussian.

Observ. Eq $\rightarrow$ $L \left\{ \vec{Y}_n = C_n \vec{X}_n + \vec{W}_n \right.$ $\searrow$ State Estimation : Kalman Filter

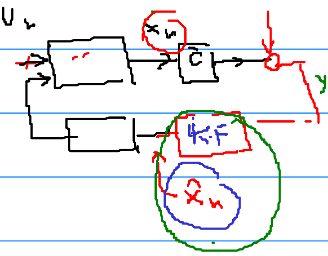
observable. $[C, CA, CA^2 \cdots CA^{k-1}]$

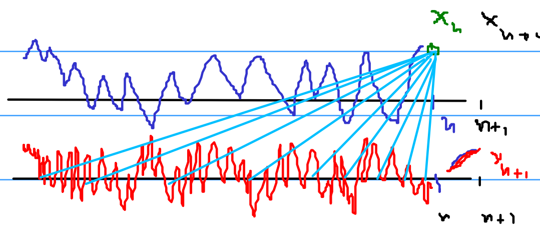
$$\vec{X}_n = f_n(\vec{X}_{n-1}, \vec{U}_n)$$

$\rightarrow$ Linearization $\rightarrow$ Extended Kalman Filters

$$Y_n = h_n(X_n, W_n)$$

$\searrow$ Partial filters





$$\hat{X}_{n|n} = \mathbb{E}[X_n \mid Y_{0:n}] = \text{Best. M.S. estimator of } X_n \mid Y_{0:n}$$

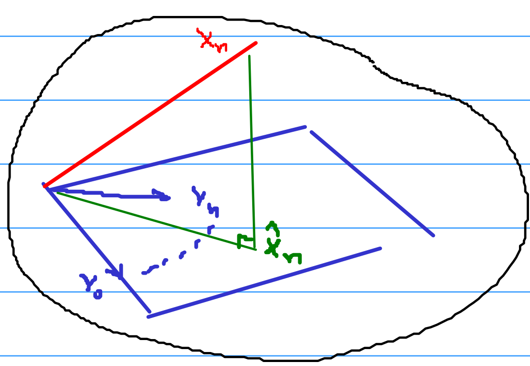
$$= \arg\min_{\phi} \mathbb{E}\left(|X_n - \phi(Y_{0:n})|^2 \right)$$

for the gaussian model

$$\hat{\vec{X}}_n = \mathbb{E}[X_n \mid \underbrace{Y_{0:n}}_{\underline{Y}_n}] = L_n \underline{Y} \quad ; \quad L_n = \underbrace{\mathbb{E}[X_n \underline{Y}_n^T] \cdot \mathbb{E}[\underline{Y}_n \underline{Y}_n^T]^{-1}}_{\text{Matrix}}$$

$\rightarrow L_{n+1}$

X_n



$$\hat{X}_{n|n} = \text{Proj}(X_n \mid Y_{0:n})$$

$$\underbrace{\hat{x}_{n|n}}_{\text{Estimated}} = \text{Proj}(x_n | y_{0:n}) = \text{Proj}(x_n | y_{0:n-1}, \tilde{y}_{n|n-1})$$

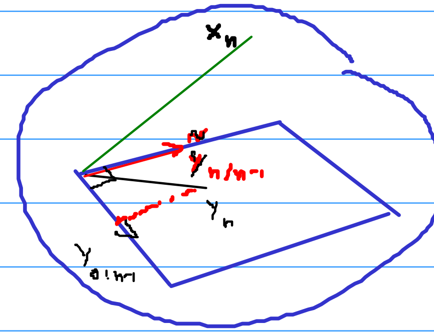
Innovation

$$= \underbrace{\text{Proj}(x_n | y_{0:n-1})}_{\hat{x}_{n|n-1}} + \underbrace{\text{Proj}(x_n | \tilde{y}_{n|n-1})}_{K_n \cdot \tilde{y}_{n|n-1}}$$

$$\tilde{y}_{n|n-1} = y_n - \text{Proj}(y_n / y_{0:n-1})$$

Predicted

gain



* $\hat{x}_{n|n-1} = \text{Proj}(x_n | y_{0:n-1})$

$$= \text{Proj}(A_n x_{n-1} + u_n | y_{0:n-1}) = A_n \overbrace{\text{Proj}(x_{n-1} | y_{0:n-1})}^{\hat{x}_{n-1/n-1}} + \underbrace{\text{Proj}(u_n | y_{0:n-1})}_{0}$$

$y_n = C_n x_n + w_n$

v_n indep $(v_0, \dots, v_{n-1})$

$$\hat{x}_{n/n} = \hat{x}_{n/n-1} + K_n \tilde{y}_n$$

* $\hat{x}_{n/n-1} = A_n \hat{x}_{n-1/n-1}$

$$\text{Proj}(w_n | y_{0:n-1}) = \cancel{E[w_n y_{0:n-1}^T]} \, E[y_{0:n-1} y_{0:n-1}^T]^{-1}$$

$\perp$

* $\tilde{y}_n = y_n - \text{Proj}(y_n | y_{0:n-1}) = y_n - \text{Proj}(C_n x_n + \overbrace{w_n}^{\perp} | y_{0:n-1}) = y_n - C_n \underbrace{\text{Proj}(x_n | y_{0:n-1})}_{\hat{x}_{n/n-1}}$

* $\tilde{y}_n = C_n x_n + w_n - C_n \hat{x}_{n/n-1} = C_n \underbrace{(x_n - \hat{x}_{n/n-1})}_{\tilde{x}_{n/n-1}} + w_n$

$\perp$

$y_n = C_n x_n + w_n$

* $K_n = \underbrace{E[x_n \tilde{y}_n^T]}_{①} \underbrace{E[\tilde{y}_n \tilde{y}_n^T]^{-1}}_{②}$

$$① = E[x_n (C_n \tilde{x}_{n/n-1} + w_n)^T] = \underbrace{E[\tilde{x}_{n/n-1} \tilde{x}_{n/n-1}^T]}_{P_{n/n-1}} C_n^T$$

$$x_n = \underbrace{\hat{x}_{n/n-1}}_{\perp} + \tilde{x}_{n/n-1}$$

$$② = \text{Cov}(\tilde{y}_n) = \text{Cov}(C_n \underbrace{\tilde{x}_{n/n-1}}_{\perp} + w_n) = C_n \underbrace{\text{Cov}(\tilde{x}_{n/n-1})}_{P_{n/n-1}} C_n^T + \underbrace{\text{Cov}(w_n)}_{R_w(n)}$$

* $P_{n/n-1} = \text{cov}(\tilde{x}_{n/n-1}) = \text{cov}(x_n - \hat{x}_{n/n-1}) = \text{Cov}(A_n x_{n-1} + u_n - A_n \hat{x}_{n-1/n-1})$

$$= \text{Cov}(A_n \underbrace{(x_{n-1} - \hat{x}_{n-1/n-1})}_{\tilde{x}_{n-1/n-1}} + u_n)$$

$\perp$

$$= \text{Cov}(A_n \tilde{x}_{n-1/n-1}) = A_n \overbrace{\text{Cov}(\tilde{x}_{n-1|n-1})}^{P_{n-1/n-1}} A_n^T$$

* $\underbrace{P_{n/n}} = \mathrm{Cov}\,(x_n - \hat{x}_{n/n}) = \mathrm{cov}\,(x_n - [\hat{x}_{n/n-1} + K_n \tilde{y}_{n}])$

$\tilde{y}_n = c_n \tilde{x}_{n/n-1} + w_n$

$= \mathrm{cov}\left[(I - K_n c_n)\, \underbrace{\tilde{x}_{n/n-1} - K_n w_n}_{\perp} \right]$

$= (I - K_n c_n)\, \underbrace{\mathrm{Cov}(\tilde{x}_{n/n-1})}_{P_{n/n-1}}\, (I - K_n c_n)^{T} + K_n \underbrace{\mathrm{Cov}(w_n)}_{R_w(n)} K_n^{T}$

$\tilde{x}_{n/n-1} = x_n - \hat{x}_{n/n-1}$

$= x_n - A_n \hat{x}_{n-1/n-1}$