Sufficient Statistic

Def: U, V, W are said to form a Markov chain, denoted $U-V$

iff $\quad p_{W|U,V}(w|u,v) = p_{W|U}(w|u)$

or equivalently $\quad p_{U|V,W}(u|v,w) = p_{U|V}(u|v)$

Def: We want to infer X.

We observe $\bar{y}$.

Let T be a r.v obtained from processing $\bar{y}$

i.e $\quad T = h(\bar{y})$.

T is a sufficient statistic if $X - T - \bar{y}$

Note: $\quad X - \bar{y} - T$ is always true if $T = h(\bar{y})$

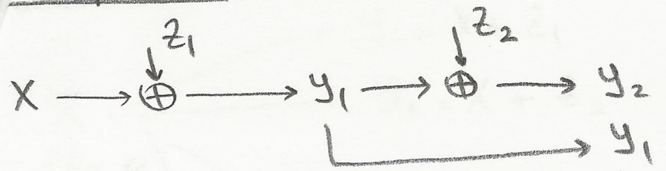
$$p_{T|\bar{y},X}(t|\bar{y},x) = \mathbb{1}\{t = h(\bar{y})\} = p_{T|\bar{y}}(t|\bar{y})$$

If T is sufficient statistic, $X - T - \bar{y}$

$$p_{X|\bar{y}}(x|\bar{y}) \underset{\substack{\uparrow \\ \text{always} \\ \text{true}}}{=} p_{X|\bar{y},T}(x|\bar{y},t) \underset{\substack{\uparrow \\ \text{if } T \text{ is a} \\ \text{sufficient statistic}}}{=} p_{X|T}(x|t)$$

$\Rightarrow$ T is sufficient to optimally decode / estimate X

xample:



Z_1 and Z_2 are independent and also independent of X.

$T = Y_1$ is a sufficient statistic.

$$P_{X|Y_1,Y_2}(x|y_1,y_2) = P(X=x \mid Y_1=y_1, Y_2=y_2)$$
$$= P(X=x \mid Y_1=y_1,\ Y_1+Z_2=y_2)$$
$$= P(X=x \mid Y_1=y_1,\ Z_2=y_2-y_1)$$
$$= \frac{P(X=x,\ Y_1=y_1,\ Z_2=y_2-y_1)}{P(Y_1=y_1,\ Z_2=y_2-y_1)}$$
$$= \frac{P(X=x,\ Y_1=y_1)\,\cancel{P(Z_2=y_2-y_1)}}{P(Y_1=y_1)\,\cancel{P(Z_2=y_2-y_1)}}$$
$$= P(X=x \mid Y_1=y_1)$$

Equivalently

$$P_{Y_2|Y_1,X}(y_2|y_1,x) = P_{Y_2|Y_1}(y_2|y_1)$$

$$Y_2 - Y_1 - X$$

$$Y_2 = \boxed{Y_1} + Z_2$$

$$Y_1 = \boxed{X} + Z_1$$

$\Rightarrow X - Y_1 - Y_2$, $X - Y_1 - (Y_1, Y_2)$

Y_1 is a sufficient statistic, Y_2 is irrelevant.

Example 2: $Y_1 = X + Z_1$
$Y_2 = X + Z_1 + Z_2$

Z_1 and Z_2 are independent of X but correlated with each other

e.g $Z_1 = Z_2 \Rightarrow Y_2 - Y_1 = Z_2$ $\quad Y_1 - Z_2 = X$.

$\Rightarrow Y_1$ is not sufficient statistic.

Example 3: $Y_1 = X + Z_1$
$Y_2 = X + Z_2$

Z_1 and Z_2 independent.

Example 4: $Y_1 = h_1 X + Z_1$ $\quad Y_2 = Y_1/h_1 = X + Z_1/h_1$.

Recursive Estimation:

$X \sim N(0,P)$

- $y_1 = X + Z_1 \qquad Z_1 \sim N(0, \sigma_1^2)$
- $y_2 = X + Z_2 \qquad Z_2 \sim N(0, \sigma_2^2)$
- $y_n = X + Z_n \qquad Z_n \sim N(0, \sigma_n^2)$

$Z_1, Z_2, \ldots Z_n$ are independent of each other and X.

Last time: $\mathbb{E}[X \mid y_1, \ldots y_n]$.

Recursive Estimation

$\underline{n=1:}$

$$X \sim N(0, P)$$

$$y_1 = X + Z_1 \qquad Z_1 \sim N(0, \sigma_1^2)$$

$$\hat{X}_1 = \frac{P}{P+\sigma_1^2} y_1 = k_1 y_1$$

$$e_1^2 = MSE = \frac{P\sigma_1^2}{P+\sigma_1^2} = k_1 \sigma_1^2$$

$\underline{n=2}$

$$\hat{X}_1, \; \cancel{y_1 = X + Z_1} \qquad y_2 = X + Z_2$$

$$X = \boxed{\hat{X}_1} + W_1 \qquad W_1 \sim N(0, \sigma_w^2 = e_1^2)$$

$$\boxed{y_2} = X + Z_2 = \boxed{\hat{X}_1} + W_1 + Z_2$$

$\boxed{y_1}$ is irrelevant.

Given $\hat{X}_1$, $X \sim N(\hat{X}_1, e_1^2)$

$$y_2 = X + Z_2$$

$$X_2 = \hat{X}_1 + \frac{e_1^2}{e_1^2+\sigma_2^2}(y_2 - \hat{X}_1) = \underbrace{\frac{\sigma_2^2}{e_1^2+\sigma_2^2}}_{1-k_2} \hat{X}_1 + \underbrace{\frac{e_1^2}{e_1^2+\sigma_2^2}}_{k_2} y$$

$$MSE = \frac{e_1 \sigma_2^2}{e_1 + \sigma_2^2} = \sigma_2^2 k_2 = e_2^2$$

$n = 3$

$$y_3 = X + z_3 \qquad \hat{X}_2 \, , \, y_2$$

$$X = \hat{X}_2 + W_2$$

$$X \mid \hat{X}_2 = \hat{x}_2 \sim N(\hat{x}_2, e_2^2)$$

$$\hat{X}_3 = \hat{X}_2 + \frac{e_2^2}{e_2^2 + \sigma_3^2}(y_3 - \hat{X}_2) = \underbrace{\frac{\sigma_3^2}{e_2^2 + \sigma_3^2}}_{1-k_3} \hat{X}_2 + \underbrace{\frac{e_2^2}{e_2^2 + \sigma_3^2}}_{k_3} y_3$$

$$= (1 - k_3)\hat{X}_2 + k_3 y_3$$

$$e_3 = MSE = \frac{\sigma_3^2 e_2^2}{e_2^2 + \sigma_3^2} = k_3 \sigma_3^2$$

At step n:

$$\hat{X} = k_n y_n + (1 - k_n)\hat{X}_{n-1}$$

$$k_n = \frac{e_{n-1}^2}{e_{n-1}^2 + \sigma_n^2}$$

$$e_n^2 = k_n \sigma_n^2$$

Take e_{n-1} and $\hat{X}_{n-1}$ from step $n-1$ and apply the above recursion.

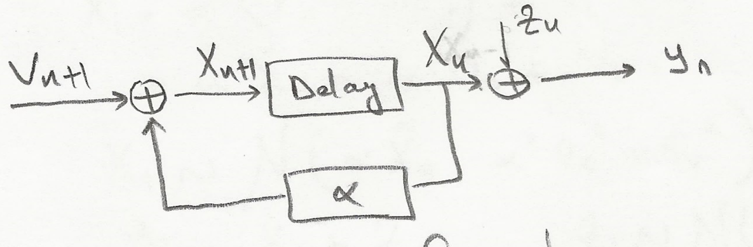
Example: Kalman Filtering

Linear Dynamical System

Initial State : $X_0 \sim N(0, \sigma_0^2)$

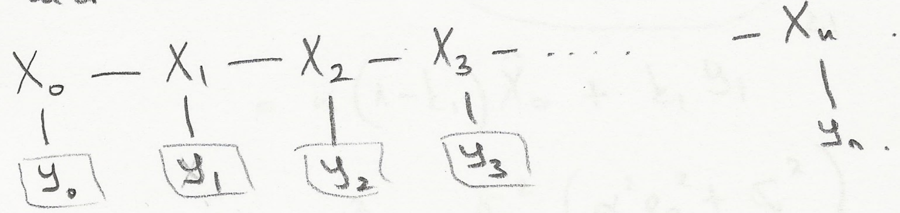
State at time n: $X_n = \alpha X_{n-1} + V_n \quad V_n \sim N(0, \sigma_v^2)$

Observation : $Y_n = X_n + Z_n$ where $Z_n \sim N(0, \sigma_z^2)$



Examples: Navigation, face tracing, stock market, robotics.

Important: V's and Z's are independent from each other and X_0



Special case : $\sigma_v^2 = 0 . \quad V_n = 0 . \quad \alpha = 1 .$

$$X_n = X_{n-1} = X_{n-2} = \dots = X_0 .$$

n = 0

$$X_0 \sim N(0, \sigma_0^2)$$

$$y_0 = X_0 + Z_0 .$$

$$\hat{X}_0 = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_z^2} y_0 \qquad e_0^2 = \frac{\sigma_0^2 \sigma_z^2}{\sigma_0^2 + \sigma_z^2}$$

$\underline{n=1:}$ $\rightarrow X_1 = \alpha X_0 + V_1$.

$y_1 = X_1 + Z_1$.

$\hat{X}_0 \ , \ \cancel{y_0}$ $\quad X_0 = \hat{X}_0 + W_0$. $W_0 \sim N(0, e_0^2)$

$W_0 \perp y_0$.

$$X_1 = \alpha(\hat{X}_0 + W_0) + V_1$$

$$= \alpha \hat{X}_0 + \alpha W_0 + V_1$$

$$X_1 \sim N(\alpha \hat{X}_0 , \alpha^2 e_0^2 + \sigma_v^2)$$

$$y_1 = X_1 + Z_1 \qquad Z_1 \sim N(0, \sigma_z^2)$$

$$\hat{X}_1 = \alpha \hat{X}_0 + \underbrace{\frac{\alpha^2 e_0^2 + \sigma_v^2}{\alpha^2 e_0^2 + \sigma_v^2 + \sigma_z^2}}_{k_1} (y_1 - \alpha \hat{X}_0)$$

$$= \alpha(1-k_1)\hat{X}_0 + k_1 y_1$$

$$\text{MSE} : e_1^2 = \frac{\sigma_z^2(\alpha^2 e_0^2 + \sigma_v^2)}{\alpha^2 e_0^2 + \sigma_v^2 + \sigma_z^2} = k_1 \sigma_z^2$$

$\underline{n=3}$

$$X_2 = \alpha X_1 + V_2$$

$$y_2 = X_2 + Z_2$$

$\hat{X}_1 \ , \ \cancel{y_1}$ $\quad X_1 = \alpha \hat{X}_1 + W_1$ $\quad W \sim N(0, e_1^2)$

$W \perp y_1$

$$X_2 = \alpha \hat{X}_1 + \alpha W_1 + V_2$$

$$X_2 \sim N(\alpha \hat{X}_1 , \alpha^2 e_1^2 + \sigma_v^2)$$

$$\hat{X}_2 = \alpha \hat{X}_1 + \frac{\alpha^2 e_1^2 + \sigma_v^2}{\alpha^2 e_1^2 + \sigma_v^2 + \sigma_z^2} (y_2 - \alpha \hat{X}_1) \longrightarrow k_2$$

$$= \alpha(1-k_2)\hat{X}_1 + k_2 y_2$$

$e_2 = k_2 \sigma_z^2$

At time step n: We have $\hat{X}_{n-1}$ and e_{n-1}^2 from the previous step

$$k_n = \frac{\alpha^2 e_{n-1}^2 + \sigma_v^2}{\alpha^2 e_{n-1}^2 + \sigma_v^2 + \sigma_z^2}$$

$$\hat{X}_n = \alpha(1-k_n)\hat{X}_{n-1} + k_n y_n$$

$$e_n^2 = k_n \sigma_z^2$$