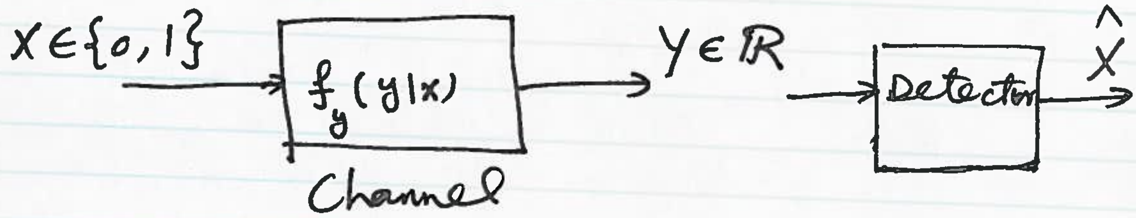
X Lecture. 6, Oct. 13, 2004

An example of using the Bayes Rule

Assume that a transmitter sends two symbols 0 and 1 and the receiver receives a symbol y with the conditional probability density $f_Y(y|X=x)$, $x=0,1$



Assume we have received $Y=y$, we can decide either that X was equal to zero or X was equal to one.

Assume, we decide $\hat{X}=0$, what is the probability of error $P_{\varepsilon,0} = P(X=1|Y=y) = P(X=1|y)$

$$= 1 - P(X=0|y)$$

but if we decide $\hat{X}=1$,

$$P_{\varepsilon,1} = P(X=0|y) = 1 - P(X=1|y)$$

If we are interested in reducing the probability of error, we have to decide in favor of the choice with lower probability

So, the decision rule will be:

decide $\hat{X} = 0$ if $P(X=0|y) \geq P(X=1|y)$

decide $\hat{X} = 1$ if $P(X=1|y) > P(X=0|y)$

This is called the <u>Maximum</u> <u>a</u> <u>posterior</u> Probability (MAP) detection.

<u>Binary Communication</u> over AWGN channel:

Assume that bits are encoded as

$$0 \rightarrow -A \text{ volts}$$
$$1 \rightarrow +A \text{ ''}$$

and transmitted over Additive White Gaussian Noise (AWGN) channel. Then

$$f_Y(y|0) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y+A)^2}{2\sigma^2}}$$

and

$$f_Y(y|1) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-A)^2}{2\sigma^2}}$$

if probability of send a zero is $P(0) = p_0$ and probability of sending a one is $P(1) = p_1 = 1 - p_0$, then

$$P(X=0|y) = \frac{P(X=0) f_Y(y|0)}{f_Y(y)} = \frac{p_0 f_Y(y|0)}{f_Y(y)}$$

and

$$P(X=1|y) = \frac{P(X=1) f_Y(y|1)}{f_Y(y)} = \frac{p_1 f_Y(y|1)}{f_Y(y)}$$

Based on the MAP detection criterion, the decision rule would be:

$$\text{Decide } \hat{X} = 0 \text{ if } \frac{P_0 f_Y(y|0)}{f_Y(y)} \geq \frac{P_1 f_Y(y|1)}{f_Y(y)}$$

and

$$\text{decide } \hat{x} = 1 \text{ if } \frac{P_0 f_Y(y|0)}{f_Y(y)} < \frac{P_1 f_Y(y|1)}{f_Y(y)}$$

Since $f_Y(y)$ is common to both sides, we have

$$\text{decide } \hat{x} = 0 \text{ if } \frac{f_Y(y|1)}{f_Y(y|0)} \leq \frac{P_0}{P_1}$$

and

$$\text{decide } \hat{x} = 1 \text{ if } \frac{f_Y(y|1)}{f_Y(y|0)} > \frac{P_0}{P_1}$$

$\frac{f_Y(y|1)}{f_Y(y|0)}$ is called the Likelihood Ratio

Substituting for $f_Y(y|1)$ and $f_Y(y|0)$ we get

$$\text{decide } \hat{x} = 1 \text{ if } e^{-\frac{(y-A)^2}{2\sigma^2} + \frac{(y+A)^2}{2\sigma^2}} > \frac{P_0}{P_1}$$

Taking the logarithm of both sides, we get:

$$-\frac{(y-A)^2}{2\sigma^2} + \frac{(y+A)^2}{2\sigma^2} > \ln \frac{P_0}{P_1}$$

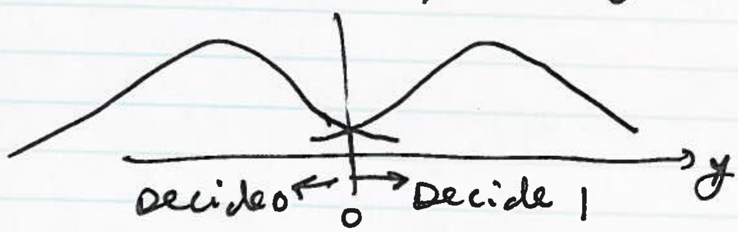
when the bits are equiprobable, i.e., when $P_1 = P_0 = \frac{1}{2}$

we have $\ln \frac{P_0}{P_1} = 0$ and we have

Decide $\hat{x}=1$ if $(y-A)^2 < (y+A)^2$ decide $\hat{x}=0$ otherw

That is, decide based on the nearest-neighbor criteri

$$\cancel{y^2}+\cancel{A^2}-2Ay<\cancel{y^2}+\cancel{A^2}+2Ay \Rightarrow y>0$$



Decide $\hat{x}=1$ if $y>0$ decide $\hat{x}=0$ otherwise

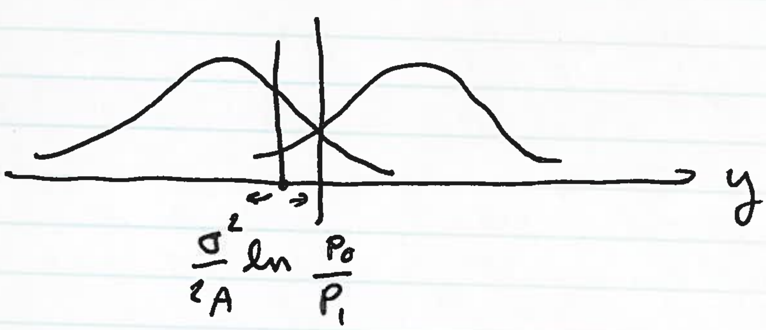
when $P_0 \neq P_1$ we have:

$$-(y-A)^2+(y+A)^2 > 2\sigma^2 \ln \frac{P_0}{P_1} \Rightarrow \hat{x}=1$$

~~simplifying the~~

Simplifying this, we get:

Decide $\hat{x}=1$ if $y > \frac{\sigma^2}{2A} \ln \frac{P_0}{P_1}$



Probability of error

Take the case of $P_0 = P_1 = \frac{1}{2}$.

Then there are two types of errors:

1) $X=0$, but $\hat{X}=1$, i.e., $y>0$

2) $X=1$, but $\hat{X}=0$, i.e., $y<0$

So; the probability of error is:

$$P_\varepsilon = P_0\, P(y>0 \mid X=0) + P_1\, P(y<0 \mid X=1)$$

$$= \frac{1}{2}\cdot\frac{1}{\sqrt{2\pi}\sigma}\int_0^\infty e^{-\frac{(y+A)^2}{2\sigma^2}}dy + \frac{1}{2\sqrt{2\pi}\sigma}\int_{-\infty}^0 e^{-\frac{(y-A)^2}{2\sigma^2}}dy$$

Due to Symmetry, both integrals can be shown to be equal so,

$$P_\varepsilon = \frac{1}{\sqrt{2\pi}\sigma}\int_0^\infty e^{-\frac{(y+A)^2}{2\sigma^2}}dy$$

let $u = \frac{y+A}{\sigma}$, then

$$\therefore P_\varepsilon = \frac{1}{\sqrt{2\pi}}\int_{\frac{A}{\sigma}}^\infty e^{-u^2} du = Q\left(\frac{A}{\sigma}\right)$$

Conditional Expectation

$$E[Y|x] = \int_{-\infty}^{\infty} y\, f_y(y|x)\, dy$$

and in discrete case

$$E[Y|x] = \sum_{y_i} y_i\, P_y(y_i|x)$$

B) Multiple Random Variables

$$F_{X_1,\dots X_n}(x_1,\dots x_n) = P[X_1 \le x_1, \dots, X_n \le x_n]$$

$$f_{X_1\dots X_n}(x_1,\dots x_n) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n} F_{X_1 \dots X_n}(x_1,\dots x_n)$$

marginal pdf

$$f_{X_i}(x_i) = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{\text{all } x_1 \dots x_n \text{ except } x_i} f(x_1,\dots x_n)\, dx_1 \cdots dx_n$$

example:

$$f_{X_1 X_2 X_3}(x_1 x_2 x_3) = \frac{e^{-(x_1^2 + x_2^2 - \sqrt{2}x_1 x_2 + \frac{1}{2}x_3^2)}}{2\pi\sqrt{\pi}}$$

$$f_{X_1,X_3}(x_1,x_3) = \frac{e^{-x_3^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{-(x_1^2+x_2^2-\sqrt{2}\,x_1 x_2)}}{2\pi/\sqrt{2}}\, dx_2$$

or

$$f_{X_1 X_3}(x_1,x_3) = \frac{e^{-x_3^2/2}}{\sqrt{2\pi}} \cdot \frac{e^{-x_1^2/2}}{\sqrt{2\pi}}$$

Function of several random variables

$$Z = g(X_1, X_2, \ldots X_n)$$

First find

$$R_z = \{(x_1, x_2, \ldots x_n) = \underline{x} \text{ such that } g(\underline{x}) \le z\}$$

then

$$F_Z(z) = P[\underline{X} \text{ in } R_z]$$

$$= \int_{\underline{x} \in R_z} \cdots \int f_{X_1 \ldots X_n}(x_1' \ldots x_n')\, dx_1' \cdots dx_n'$$

Example

$Z = X + Y.$ $\qquad Z \le z \Rightarrow X + Y \le z$

$\qquad\qquad\qquad\qquad Y \le z - x$

Find $f_Z(z)$ in terms of $f_X(x)$ and $f_Y(y)$

$$F_Z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x'} f_{X,Y}(x', y')\, dy'\, dx'$$

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x', z - x')\, dx'$$

if X and Y are independent

then

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x') f_Y(z - x')\, dx'$$

$$\cancel{= f_X(x) * f_Y(y)}$$

Example: X and Y are two independent random variables with exponential distribution with mean one.

Find the pdf of $Z = \frac{X}{Y}$

let $Y = y$ then $Z = \frac{X}{y}$ (a scaled version of X)

$f_Z(z|y) = |y| f_X(yz|y)$ // since $f_X(ax) = \frac{1}{|a|} f_X(\frac{x}{a})$

$$f_Z(z) = \int_{-\infty}^{\infty} |y| f_X(yz|y) f_Y(y)\, dy$$

$$f_Z(z) = \int_{-\infty}^{\infty} |y|\, f_X(yz)\, f_Y(y)\, dy$$

$$f_Z(z) = \int_0^{\infty} y\, e^{-yz} \cdot e^{-y}\, dy = \frac{1}{(z+1)^2} \quad z > 0$$

Transformations of Random Vectors

Let $X_1, \cdots X_n$ be random variables and let

$Z_1 = g_1(X_1, \cdots X_n)$, $Z_2 = g_2(X_1, \cdots X_n) \cdots Z_n = g_n(X_1, \cdots X_n)$

The joint CDF of $\underline{Z}$ is given as

$$F_{Z_1 \cdots Z_n}(z_1, \cdots z_n) = P[g_1(\underline{X}) \le z_1 \cdots g_n(\underline{X}) \le z_n]$$

$$= \int \cdots \int_{\underline{x} : g(\underline{x}) \le z_k} f_{X_1 \cdots X_n}(x_1' \cdots x_n')\, dx_1' \cdots dx_n'$$

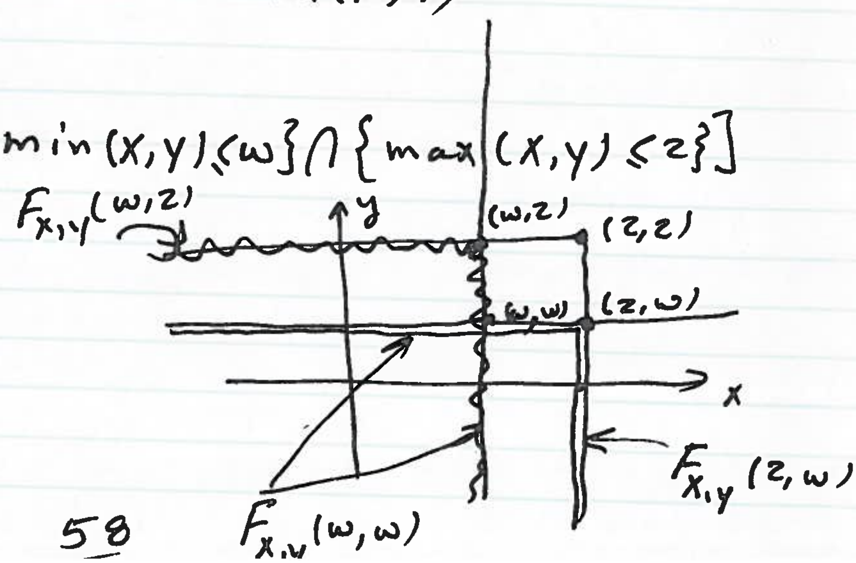
Example: Let

$$W = \min(X, Y) \qquad Z = \max(X, Y)$$

then

$$F_{W,Z}(w, z) = P[\{\min(X,Y) \le w\} \cap \{\max(X,Y) \le z\}]$$

$$F_{W,Z}(w,z) = F_{X,Y}(z,w) + F_{X,Y}(w,z) - F_{X,Y}(w,w)$$



pdf of linear transformation

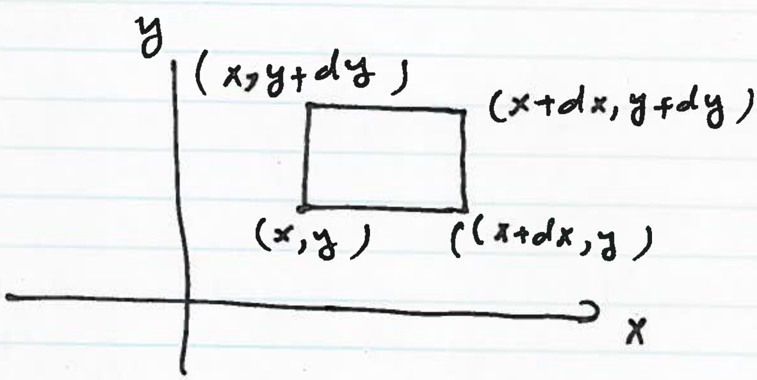
$$V = aX + bY \qquad \Rightarrow \qquad \begin{bmatrix} V \\ W \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}$$
$$W = cX + dY$$

assume that A is not Singular ($ad \neq bc$)

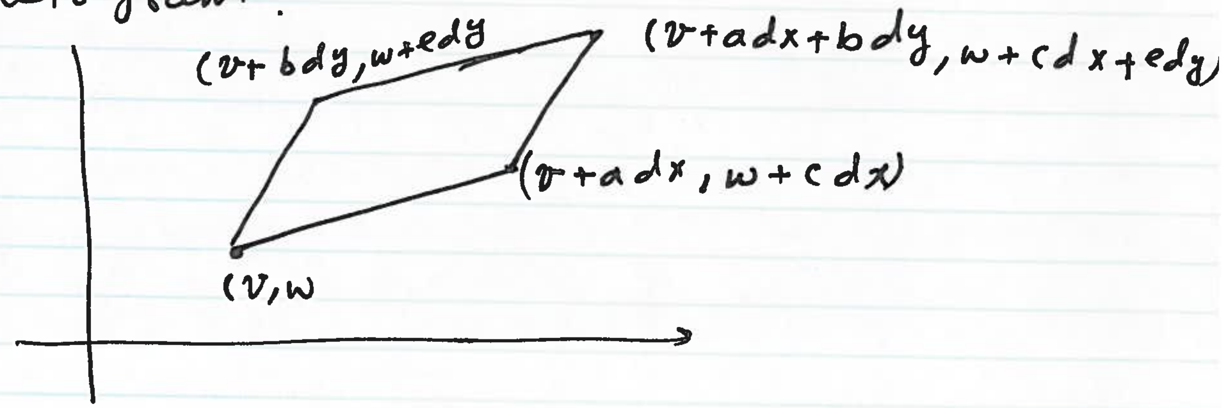
then for each (v, w) pair, we have a unique (x, y) pair:

$$\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1} \begin{bmatrix} v \\ w \end{bmatrix}$$

Consider infinitesimal rectangle:



This rectangle will be transformed into ~~the~~ parallelogram:



The rectangle and the parallelogram represent equivalent events, so their probabilities should be equal

$$f_{x,y}(x,y)\,dx\,dy = f_{V,W}(v,w)\,dP$$

So:

$$f_{V,W}(v,w) = \frac{f_{x,y}(x,y)}{\left|\frac{dP}{dx\,dy}\right|}$$

it can be shown that

$$dP = (|ae-bc|)\,dx\,dy$$

So

$$\frac{dP}{dx\,dy} = |ae-bc| = |A| = \det(A)$$

So

$$f_{V,W}(v,w) = \frac{f_{x,y}(x,y)}{|A|}$$ ← here (x,y) are given as $A^{-1}(v,w)$

In general if $\underline{Z}$ is a linear transform of $\underline{X} = X_1, \dots X_n$, i.e.,

$$\underline{Z} = A\underline{X}$$

we have:

$$f_{\underline{Z}}(\underline{z}) = \frac{f_{\underline{X}}(A^{-1}\underline{z})}{|A|} = \left.\frac{f_{X_1,\dots X_n}(x_1,\dots x_n)}{|A|}\right|_{\underline{x}=A^{-1}\underline{z}}$$

In general if

$$Z_1 = g_1(\underline{X}), \quad Z_2 = g_2(\underline{X}) \quad \cdots \quad Z_n = g_n(\underline{X})$$

and if the set of equations

$$z_i = g_i(x) \qquad i = 1, \ldots, n$$

have a unique solution

$$x_1 = h_1(\underline{z}) \quad x_2 = h_2(\underline{z}) \quad \cdots \quad x_n = h_2(\underline{z})$$

then

$$f_{Z_1 \cdots Z_n}(z_1 \cdots z_n) = \frac{f_{X_1 \cdots X_n}(h_1(\underline{z}), \ldots h_n(\underline{z}))}{|J(x_1 \cdots x_n)|} = f_{X_1 \cdots X_n}(h_1(\underline{z}) \cdots h_n(\underline{z})) \times |J(z_1 \cdots z_n)|$$

where the ~~denominator~~ $|J|$ is the determinant of the Jacobian matrix

$$J(x_1, \cdots x_n) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{bmatrix}$$

and

$$J(z_1 \cdots z_n) = \begin{bmatrix} \frac{\partial h_1}{\partial z_1} & \cdots & \frac{\partial h_1}{\partial z_n} \\ \vdots & & \vdots \\ \frac{\partial h_n}{\partial z_1} & \cdots & \frac{\partial h_n}{\partial z_n} \end{bmatrix}$$

Example: Let X and Y be unit-variance, zero-mean Gaussian random variables and

$$V = (X^2 + Y^2)^{1/2}$$

$$W = \angle(X, Y) = \operatorname{atan}\left(\frac{Y}{X}\right)$$

we have

$$x = \underbrace{v \cos w}_{h_1(w)} \qquad y = \underbrace{v \sin w}_{h_2(w)}$$

$$J(v, w) = \begin{bmatrix} \cos w & -v \sin w \\ \sin w & v \cos w \end{bmatrix}$$

$$|J(v, w)| = v \cos^2 w + v \sin^2 w = v$$

$$f_{V,W}(v,w) = \frac{v}{2\pi} e^{-[v^2 \cos^2 w + v^2 \sin^2 w]/2}$$

$$f_{V,W}(v, w) = \frac{1}{2\pi} v e^{-v^2/2} \qquad v \geq 0, \quad 0 \leq w \leq 2\pi$$

$$f_V(v) = \int_{-\infty}^{\infty} f_{V,W}(v, w)\, dw = \int_0^{2\pi} \frac{1}{2\pi} v e^{-v^2/2} dw = v e^{-v^2/2}$$

$$v > 0$$

this is the Rayleigh Distributed r.v.