# Random Variables and Random Process/Digital Communications/Linear Algebra Review

(6hrs)



## Review of Random Variable (R.V.)

- A probability space is a collection $(\Omega, F, P)$.
  $\Omega$: sample space containing all possible outcomes.
  F : event space, subspace of $\Omega$ with properties.
  P : probability measure.
  Random variable is a function $\Omega \to R$
  ( outcomes mapping to real numbers ).
- Data channel noise interference are all R.V.s

# Properties of Prob.

- $P(\phi)=0, \quad P(\Omega)=1$ .
- $0 \le P(A) \le 1, \quad \forall \quad A \in \Omega$ .
- *If* $A_1, A_2, A_3, \cdots\cdots$ are disjoint, then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ .

3

# Properties of R.V.

- Expectation is linear.
- $\mathrm{var}(ax+b) = a^2 \, \mathrm{var}(x)$ .
- $\mathrm{cov}(x,y) = E[xy] - E[x]E[y]$ .
- Covariance=0=> uncorrelated

4

## Discrete R.V.

- $\Omega \to 1,2,3,\cdots$
- Ex. Binomial distribution

  $P(n=k)=\binom{n}{k}P^k(1-p)^{n-k}$

  $k$ = 0,1,2,…,n ; n : the number of trials.

  Bernoulli : special case, n=1.
- Ex. Poisson distribution

  $P(X=k)=e^{-\lambda}\cdot\frac{\lambda^k}{k!},\quad k=0,1,\cdots,\infty$

5

## Continuous R.V.

- Ex. Exponential distribution

  $f(x)=1/\lambda\cdot e^{-x/\lambda}$
- Ex. Gaussian distribution

  $f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}\cdot e^{\frac{-(x-m_x)^2}{2\sigma^2}} \sim N(m_x,\sigma^2)$

6

- Ex. Multivariate Gausscian R.V.

$$f(\vec{x}) = \frac{1}{(2\pi)^{\frac{N}{2}}|R|^{\frac{1}{2}}} \cdot \exp[\tfrac{-1}{2}(\vec{x}-\vec{m})^T R^{-1}(\vec{x}-\vec{m})]$$

Covariance matrix :

$$R = E[\vec{x}\,\vec{x}^T] - \vec{m}\,\vec{m}^T$$

mean :

$$\vec{m} = E[\vec{x}]$$

7

- Ex. Chi [kai]- square distribution

$x_n^2$ with n degree of freedom.

$$x_n^2 \overset{\Delta}{=} \sum_{i=1}^{n} y_i^2; \quad y_i \sim N(0,1) \quad i.i.d.$$

$$f_{x_n^2}(x) = \frac{1}{2^{n/2}\Gamma(n/2)} \cdot x^{\frac{n}{2}-1} \cdot e^{-\frac{x}{2}}; \quad x \ge 0;$$

*where* $\Gamma(\cdot)$ *is gamma function* .

$$\Gamma(n+1) = n\Gamma(n) = \int_0^\infty y^n e^{-y} dy; \quad n > -1$$

$$\Gamma(n+1) = n! \quad \text{if } n \text{ is integer}$$

$$\Gamma(1) = 1, \quad \Gamma(1/2) = \sqrt{\pi}$$

8

# Marginal p.d.f.

- Conditional p.d.f.

$$f(x \mid y) = f(x, y) / f(y) \xrightarrow{\text{if } x,y \text{ is indep.}} f(x)$$

- Conditional expectation

$$E[x \mid y] = \int x \cdot f(x \mid y) dx;$$

$$E_y[E_x[x \mid y]] = E[x] = E[E[x \mid y]]$$



- Ex.

| coin1 | coin2 | P |
|---|---|---|
| H | H | 0 |
| H | T | 0.25 |
| T | H | 0.25 |
| T | T | 0.5 |

$$P(coin1 = H \mid coin2 = H) \stackrel{?}{=} P(coin1 = H)$$

$$\Rightarrow 0 / 0.25 \neq 0.25$$

$$\Rightarrow coin1 = H \quad and \quad coin2 = H \quad not \quad independent.$$

$$note: if \quad x, y \quad are \quad independent,$$

$$then \quad f(x \mid y) = f(x) \forall x, y$$

- Ex. Bivariate Gaussian distribution

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\cdot\exp\{\frac{-1}{2(1-r^2)}[\frac{(x-m_1)^2}{\sigma_1^2}$$

$$-2r\frac{(x-m_1)(y-m_2)}{\sigma_1\sigma_2}+\frac{(y-m_2)^2}{\sigma_2^2}]\}\sim N(m_1,m_2,\sigma_1,\sigma_2,r)$$

$$f(x\mid y)\sim N(r\frac{\sigma_1}{\sigma_2}(y-m_2)+m_1,\sigma_1\sqrt{1-r^2})$$

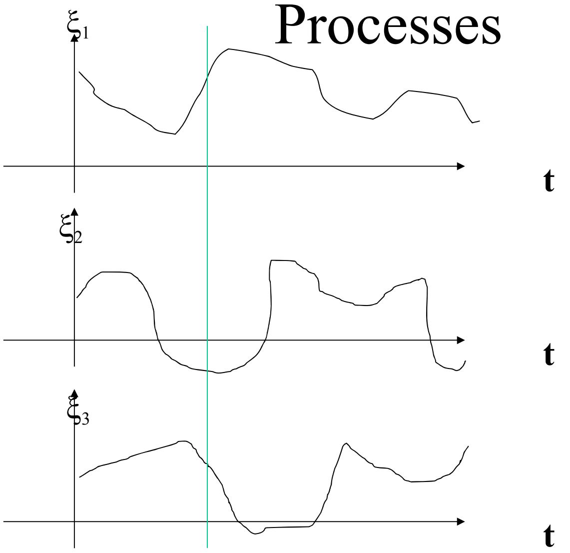
**(proof is left for the readers)**

$$\Rightarrow E[x\mid y]=r\frac{\sigma_1}{\sigma_2}(y-m_2)+m_1$$

$$R=\begin{bmatrix}\sigma_1^2 & r\sigma_1\sigma_2\\ r\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix}$$

11

# Random Variables and Random Processes




12

- Random variable (r.v): One outcome $\xi_i$ =>(maps to) one real number $x(\xi_i)$
- Random process (r.p): One outcome $\xi_i$ => one function of time $x(t,\xi_i)$
- $t,\xi$ fixed=> a real number
- t fixed, $\xi$ variable => a r.v. (In this sense, a r.p. can be thought as time-varying r.v.'s)
- t variable ,$\xi$ fixed => a function of time
- $t,\xi$ variable=> a r. p.

13

- We define mean, autocorrelation, autocovariance for r.p x (t)
- Mean $\mu_x(t)=E[x(t)]$
- Autocorrelation $R_x(t,s)=E[x(t)x(s)]$. (multiplied by, times)
- Autocovariance $C_x(t,s)=E[(x(t)- \mu_x(t)) (x(s)- \mu_x(s))]$
- $= E[x(t)x(s)]- \mu_x(t) \mu_x(s)$
- **Note: expectation is over outcome not time.**

14

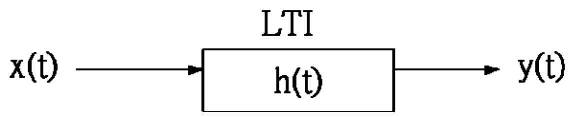
r.p. x(t) is wide - sense stationary (WSS) if

(1) $\mu_x(t) = \mu_x(0)$

(2) $R_x(t, t+\tau) = R_x(\tau)$ for all t, (2nd - 1st time index)



# Linear Filtering of Random Processes




$\tilde{h}(t) \stackrel{\Delta}{=} h(-t)$, x(t) is WSS with autocorrelation

$R_x(\tau) \stackrel{\Delta}{=} E\{x(t)x(t+\tau)\}$

$(a)\mu_y(t) = (\mu_x * h)(t)$, where $\mu_x(t) = E\{x(t)\}$

$(b)R_y(\tau) = (R_x * h * \tilde{h})(\tau)$

Proof of (a) $\mu_y(t) = E\{y(t)\} = E\{x(t) * h(t)\} = E\{\int_{-\infty}^{\infty} x(u)h(t-u)du\}$

$= \int_{-\infty}^{\infty} E\{x(u)\}h(t-u)du = \int_{-\infty}^{\infty} u_x(u)h(t-u)du = (u_x * h)(t)$

Proof of (b) $E\{y(t)y(t+\tau)\}$

$= \iint E\{x(t-\alpha)x(t+\tau-\beta)\}h(\alpha)h(\beta)d\alpha d\beta$

$= \iint R_x(\tau+\alpha-\beta)h(\alpha)h(\beta)d\alpha d\beta$

$= \iint R_x(\tau-\gamma-\beta)\widetilde{h}(\gamma)h(\beta)d\gamma d\beta$

$= (R_x * \widetilde{h} * h)(\tau)$

17

# Spectral density for WSS r.p. x(t) and y(t) Ex. Noise and filtered noise

Fourier transform pair

$R_x(\tau) \leftrightarrow S_x(f)$ (NWTT1)

$h(t) \leftrightarrow H(f)$

$\widetilde{h}(t) \leftrightarrow H^*(f)$

Thus, $R_y(\tau) = (R_x * \widetilde{h} * h)(t) \leftrightarrow S_y(f) = |H(f)|^2 S_x(f)$

18

**Important Results**

**(1) If x(t) is WSS, then y(t) is WSS.**

**(2) If x(t) is Gaussian, then y(t) is Gaussian**.

Proof of (1): from (a)and (b) above.

Proof of (2): Use the fact that linear combination of Gaussian r.v.'s is also Gaussian r.v.

($\int_{-\infty}^{\infty} g(\alpha)\cdot x(\alpha)d\alpha$ is also Gaussian)

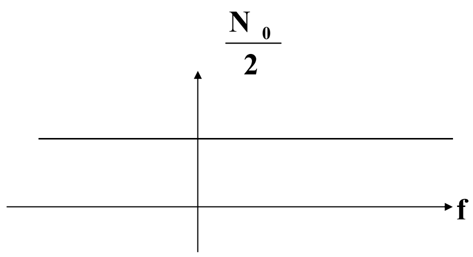
19

# Thermal Noise

Represented by additive white Gaussian noise (AWGN) random process x (t)

$S_x(f) = N_0/2$

$R_x(\tau) = \frac{N_0}{2}\delta(\tau)$




White light is composed of many colors

20

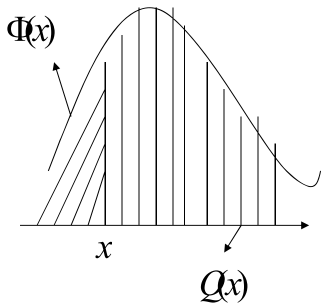
Probability density function (pdf) for N(0,1)

$$f_x(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

cumulative distribution function (cdf) for N(0,1) r.v.

$$\mathrm{F_x}(X) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = \mathrm{P_r}\{x \le x\} = \Phi(x) = 1 - Q(x)$$

$$\text{Note}: \ \Phi(\frac{\mathrm{x} - \mu}{\sigma}): \ \text{cdf of } \mathrm{N}(\mu, \sigma^2)$$




21

Comparison:

White noise: noise rvs at different times are uncorrelated

Gaussian noise: at one time, noise rv is Gaussian distributed

1)+2) noise rvs at different times are independent Gaussian

22

# Relationship of Two r.p. (optional)

Def.: crosscorrelation $R_{xy}(t,s)=E[x(t)y(s)]$

Def.: crosscovariance $C_{xy}(t,s)=E[(x(t)-\mu_x(t))(y(s)-\mu_y(s))]$

Lemma: If $R_{xy}(t,s)=\mu_x(t)\,\mu_y(s)$, or equivalently $C_{xy}(t,s)=0$, for all t and s, then x(t) and y(t) are uncorrelated.

Def.: x(t) and y(t) are independent iff for all n, $t_1,\ldots,t_n$, $s_1,\ldots,s_n$, the random vector $\underline{x}=[x(t_1),\ldots, x(t_n)]$ and $\underline{y}=[y(s_1),\ldots, y(s_n)]$ are independent. That is, x(t) and y(t) are independent iff $F_{xy}(\underline{x},\underline{y})= F_x(\underline{x})\, F_y(\underline{y})$ or equivalently $f_{xy}(\underline{x},\underline{y})= f_x(\underline{x})\, f_y(\underline{y})$.

Facts : (1) indep r.p. are uncorrelated

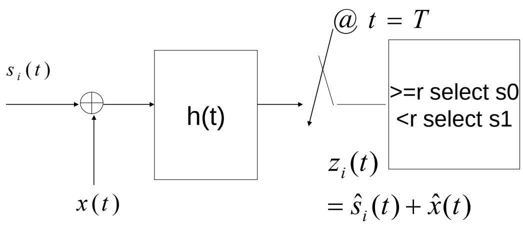
(2) jointly Gaussian+uncorrelated=>indep.

23

# Digital Communication Review

24

# General Binary Comm. System




Without loss of generality, assume $(s_0 * h)(T) >$ (is greater than) $(s_1 * h)(T)$

$z_i(t) = \int h(t-\alpha)s_i(\alpha)d\alpha + \int h(t-\alpha)x(\alpha)d\alpha = \hat{s}_i(t) + \hat{x}(t)$

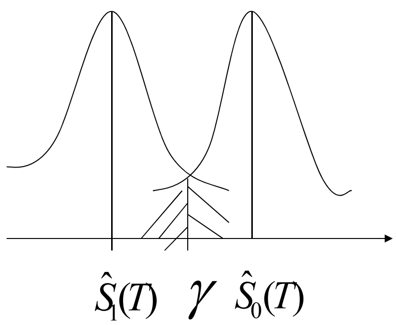


$z_i(T)$ is a Gaussian r.v. with mean $\hat{s}_i(T) = (s_i * h)(T)$

and variance $R_{\hat{x}}(0) = (h * \hat{h} * R_x)(0)$ p.16

$P(\text{error}|0\text{ sent}) = P(z_k(T) < r) = \Phi(\frac{r - \hat{s}_0(T)}{\sqrt{R_{\hat{x}}(0)}})$ p.21

$P(\text{error}|1\text{ sent}) = Q(\frac{r - \hat{s}_1(T)}{\sqrt{R_{\hat{x}}(0)}})$




Now add more constraints

(1)Assume BPSK : $s_i(t) = (-1)^i \cdot s(t)$, s(t) arbitrary

(2) Assume AWGN : Let $\tilde{h}(\mu) \overset{\Delta}{=} h(-\mu)$

$R_{\hat{x}}(\tau) = \iint R_x(\tau - \gamma - \beta)\tilde{h}(\gamma)h(\beta)\, d\gamma d\beta$

$R_{\hat{x}}(0) = \iint R_x(-\gamma - \beta)\tilde{h}(\gamma)h(\beta)\, d\gamma d\beta = \frac{N_0}{2}\int \tilde{h}(-\beta)h(\beta)d\beta$

$= \frac{N_0}{2}\int h^2(\beta)d\beta = \boxed{\frac{N_0}{2}\|h\|^2 = \frac{N_0}{2}\int |H(f)|^2 df \text{ (Parseval theorem)}}$

(3)assume $P_{e,1} = P_{e,0}$, then r = 0

# Example in a IEEE T COM paper [Wan13]

Since we assume flat fading for each subcarrier, the lowpass equivalent received signal of user $k$ on subcarrier $m$ is given by

$$y_{k,m}(t) = \sqrt{P_{k,m}} H_{k,m} X_{k,m}[l] \exp\left(\frac{j2\pi mt}{T}\right) + n_{k,m}(t) \quad (2)$$

where $n_{k,m}(t)$ is Additive White Gaussian Noise (AWGN) with two-sided power spectral density $\boxed{2N_0.}$

To detect the signal on subcarrier $m$, a correlation operation is performed $Y_{k,m} = \frac{1}{T}\int_0^T y_{k,m}(t)\exp(-j2\pi mt/T)dt$. The noise power can be calculated as $P_N = E[|N_{k,m}|^2] = \boxed{2N_0/T}$ and the power for the desired signal is $\boxed{P_{k,m}|H_{k,m}|^2.}$

p.19

$$h(t) = \begin{cases} \frac{1}{T} & 0 \le t \le T \\ 0 & otherwise \end{cases}$$

$$\text{so } \|h\|^2 = T\left(\frac{1}{T}\right)^2 = \frac{1}{T}$$



# Matched Filter (M.F.) is Optimal in AWGN

First we introduce Schwartz inequality (recall inner product in geometry in high school)

For any functions f and g on (a, b) for which

$$\|f\| \stackrel{\Delta}{=} \sqrt{\int_a^b f^2(u)du} < \infty, \ \|g\| \stackrel{\Delta}{=} \sqrt{\int_a^b g^2(u)du} < \infty,$$

$$(f,g) \stackrel{\Delta}{=} \int_a^b f(u)g(u)du, \text{ we have}$$

$(f,g)^2 \le \|g\|^2\|f\|^2$ equality holds iff there is a real number $\lambda$ such that f $= \lambda$g

Assume BPSK : $s_0(t) = s(t)$ and $s_1(t) = -s(t)$,

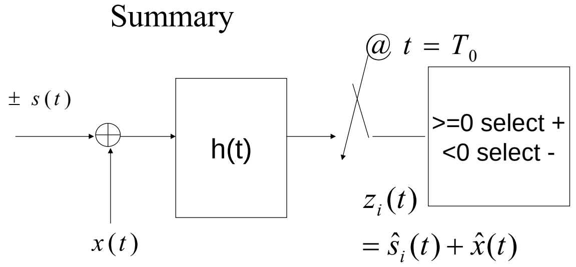
let $s_{T_0}(u) = s(\mathrm{T}-u)$, matched filter

Note $r - \hat{s}_1(\mathrm{T}) = 0 + (s * h)(\mathrm{T}) = \int \mathrm{h(u)}s(\mathrm{T}-u)du$

$$\mathrm{P_e} = Q\left(\sqrt{\frac{(\int \mathrm{h(u)}s(\mathrm{T}-u)du)^2}{\frac{\mathrm{N_0}}{2}\|h\|^2}}\right) \ge Q\left(\sqrt{\frac{\|s_{T_0}\|^2\|h\|^2}{\frac{N_0}{2}\|h\|^2}}\right) = Q\left(\sqrt{\frac{\|s_T\|^2}{\frac{N_0}{2}}}\right)$$

with equality holds iff $h(t) = \lambda s_{T_0}(t) = \lambda s(\mathrm{T}-t)$ matched filter

Summary




$$P_{e,\min} = \geq Q\left(\sqrt{\frac{\|s\|^2}{\frac{N_0}{2}}}\right) = Q\left(\sqrt{\frac{E}{\frac{N_0}{2}}}\right) = Q(\sqrt{SNR})$$

where

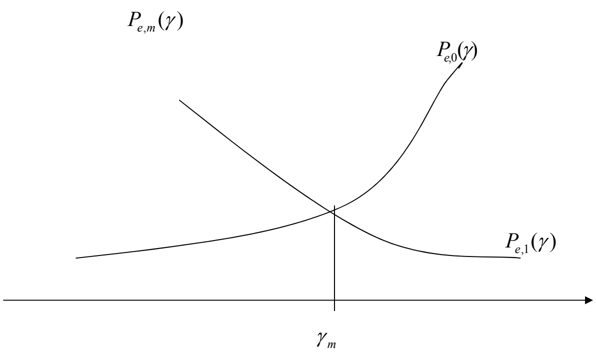
$$\boxed{\begin{array}{l}\text{received SNR} = \dfrac{E\|h\|^2}{\frac{N_0}{2}\|h\|^2} = \dfrac{E}{\frac{N_0}{2}} \text{(previous page)} \\ \text{Note : unit J} = \text{(W/(1/s))}\end{array}}$$

Q: Why Tx energy
Divided by psd =
received SNR?
A: h norm square is
ignored

29

# Minimax Approach

- minimize the maximum of $P_{e,0}$ and $P_{e,1}$ select r to give minimum of $P_{e,m}$ =max{$P_{e,0}$ and $P_{e,1}$}
- As r increases, $P_{e,0}$ increases but $P_{e,1}$ decreases.

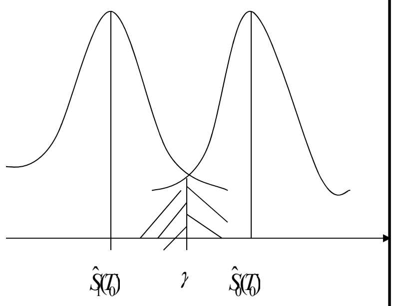



30

Choose $r = r_m$, where $r_m$ is the solution to $P_{e,0} = P_{e,1}$

$$\Phi\left(\frac{r-\hat{s}_0(T)}{\sqrt{R_{\hat{X}}(0)}}\right)=\Phi\left(\frac{\hat{s}_1(T)-r}{\sqrt{R_{\hat{X}}(0)}}\right) \Rightarrow r_m=\frac{\hat{s}_0(T)+\hat{s}_1(T)}{2}$$

$$P_{e,m}=1-\Phi\left(\frac{\frac{1}{2}(\hat{s}_0(T)-\hat{s}_1(T))}{\sqrt{R_{\hat{X}}(0)}}\right)$$



Let $s(t)\overset{\Delta}{=}s_0(t)-s_1(t)$ and $s_T(u)\overset{\Delta}{=}s(T-u)$. Then

$$P_{e,m}=1-\Phi\left(\sqrt{\frac{\frac{1}{4}(h,s_T)^2}{\frac{N_0}{2}\|h\|^2}}\right)\geq 1-\Phi\left(\sqrt{\frac{\frac{1}{4}\|s_T\|^2}{\frac{N_0}{2}}}\right)$$ with equality holds iff

$h(t)=s_T(t)=s_0(T-t)-s_1(T-t)$ matched filter

$$\|s_T(t)\|^2=\int[s_0(T-t)-s_1(T-t)]^2dt=\|s_0\|^2+\|s_1\|^2-2\int s_0(t)\cdot s_1(t)dt$$

$$=\varepsilon_0+\varepsilon_1-2(s_0,s_1)=2\bar{\varepsilon}(1-\rho)$$ where $\bar{\varepsilon}=\frac{\varepsilon_0+\varepsilon_1}{2}$, $\rho=\frac{(s_0,s_1)}{\bar{\varepsilon}}$ correlation coefficient.

$$\boxed{P_{e,m}=1-\Phi\left(\sqrt{\frac{\bar{\varepsilon}(1-\rho)}{N_0}}\right)}$$ AWGN, MF, minimax thresold

$\rho=-1$ best => BPSK best

31

# Receiver Structure

The received signal is $y(t)=s_i(t)+x(t)$

(a) M.F. implementation

(b) Correlator implementation (lower hardware cost)

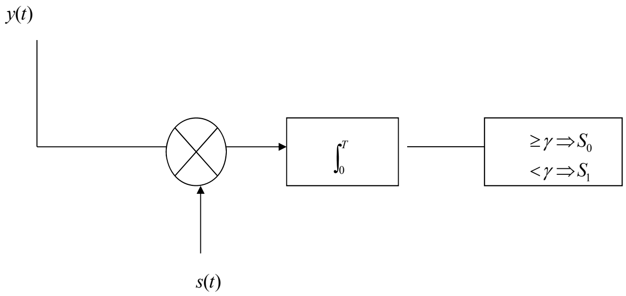
32

$$\hat{y}(T) = \int_0^T y(u)\cdot h(T-u)du$$
$$= \int_0^T y(u)\cdot s(T-(T-u))du$$
$$= \int_0^T y(u)\cdot s(u)du$$

u ~ [0, T], - u ~ [-T,0], T - u ~ [0, T]

$$\hat{y}(T) = \int_0^T y(u)\cdot s(u)du$$

when $s$ is time limited to [0, T], we have




33

# Quadriphase modulation

Transmitted signal is $R_e[Z(t)]$

where $Z(t) = \sigma(t)e^{jw_c t}$ and $\sigma(t) = [b_I(t-t_0) + jb_Q(t)]\cdot e^{j\theta}$

General form : $R_e[Z(t)] = b_I(t-t_0)\cos(W_c t+\theta) - b_Q(t)\sin(W_c t+\theta)$

$$b_I(t) = \sum_{\lambda=-\infty}^{\infty} b_\lambda^{(I)}\psi(t-\lambda T)$$

$$b_Q(t) = \sum_{\lambda=-\infty}^{\infty} b_\lambda^{(Q)}\psi(t-\lambda T) \qquad \psi(t): \text{bit waveform}$$

$$\left.\begin{array}{r} \text{BPSK} : t_0 = 0, b_Q(t) \equiv 0 \\ \text{QPSK} : t_0 = 0 \\ \text{OQPSK} : t_0 = \dfrac{T}{2} \end{array}\right\} \qquad \psi(t) = P_T(t)$$

34

# Signal Space

orthonormal set of basis function

$\phi_{\mathrm{q}}(t), q = 1,2,..........., N$ over interval $(0, \mathrm{T})$

$$\int_0^T \phi_i(t)\phi_k(t)dt = \delta_{ik}$$

$$S_i(t) = \sum_{j=1}^{N} S_{ij}\phi_j(t) \quad ,0 \le t \le T \ \ ,i = 1.......M$$

$$S_{ij} = \int_0^T S_i(t)\phi_j(t)dt = (S_i, \phi_j)$$

Generalized Fourier transform/series coefficients

Digital communication problem => Geomerty problem

35

# Gram Schmit procedure

find orthonormal basis

1) choose $\phi_1(t) = \dfrac{S_1(t)}{\| S_1 \|}$

2) $V_2 = S_2 - (S_2, \phi_1)\cdot\phi_1(t) \qquad \phi_2(t) = \dfrac{V_2(t)}{\| V_2 \|}$

3) $V_j(\mathrm{t}) = S_j - (S_j, \phi_1)\cdot\phi_1(t) \ ............ (S_j, \phi_{j-1})\cdot\phi_{j-1}(t)$

$$\phi_j(t) = \frac{V_j(t)}{\| V_j \|}$$

36

# Geometry Approach for Computing Pe

use $S_{ij}$ to represent energy , correlation, error prob

(a) Signal energy

$$\varepsilon_i = \int_0^T S_i^2(t)dt = \int_0^T \sum_{j=1}^{N}\sum_{k=1}^{N} S_{ij}S_{ik}\phi_j(t)\phi_k(t)dt = \sum_{j=1}^{N} S_{ij}^2 = \| S_i \|^2 \quad \text{where } S_i = [S_{i1},.........., S_{iN}]$$

(b) correlation coefficient

$$\rho_{ij} \overset{\Delta}{=} \frac{(S_i, S_j)}{\bar{\varepsilon}}, \text{ where } \bar{\varepsilon} \overset{\Delta}{=} \frac{\varepsilon_0 + \varepsilon_1}{2}$$

(c) minimax error prob.

$$d_{ij}^2 = \int_0^T [S_i(t) - S_j(t)]^2 dt = \sum_{k=1}^{N} (S_{ik}^2 + S_{jk}^2 - 2S_{ik}S_{jk})$$

$$= \varepsilon_i + \varepsilon_j - 2(S_i, S_j) = 2\bar{\varepsilon}(1 - \rho_{ij})$$

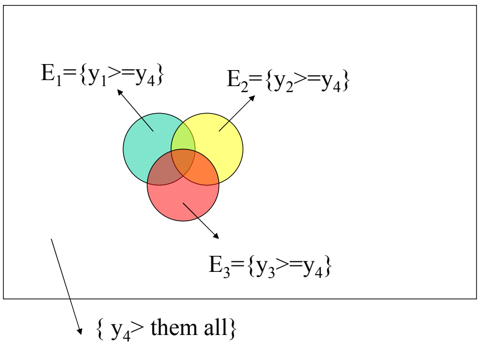
$$\therefore P_{ij} = Q(\frac{\sqrt{\bar{\varepsilon}(1-\rho)}}{\sqrt{N_0}}) = Q(\frac{d_{ij}/2}{\sqrt{N_0/2}}) \quad \text{symbol error prob. } i \Leftrightarrow j \text{ depend on distance in constellation}$$

(pp. 32) BPSK is best

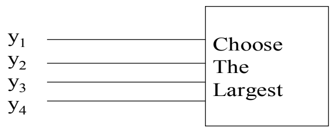
37

Union Bound ( good for high $E_b/N_0$ ) on BER

SNR=0, three circles totally overlap
SNR=infinity, three circles become three dots




38




M=4 branches
Suppose symbol 4 is sent

$$\text{Proba bility Error} = P(E_1 \cup E_2 \cup E_3)$$
$$= P(E_1) + P(E_2) + P(E_3) - P(E_1E_2) - P(E_2E_3) - P(E_3E_1) + P(E_1E_2E_3)$$
$$\le P(E_1) + P(E_2) + P(E_3)$$



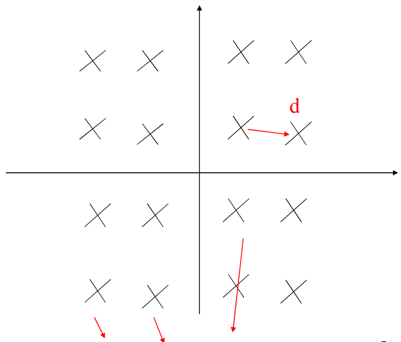
# Approximate union bound

$$\text{Prob. of symbol error} \approx N_{NB} Q\left(\frac{d/2}{\sqrt{N_0/2}}\right)$$

d $\uparrow$, Q function $\downarrow$, so we ignore all but the largest terms

$N_{NB}$ : number of nearest neighbors

# Example: 16 Quadrature Amplitude Modulation(QAM)




4/16(2Q)+8/16(3Q)+4/16(4Q) where $Q = Q(\frac{d/2}{\sqrt{N_0/2}})$



# Bit and symbol error prob.

Pr *oblem* :

How to relate bit error prob. to symbol error prob.

M - ary signals (symbols)

ex :

| | | |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 0 | 1 |
| 0 | 1 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 0 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |
| 1 | 1 | 1 |

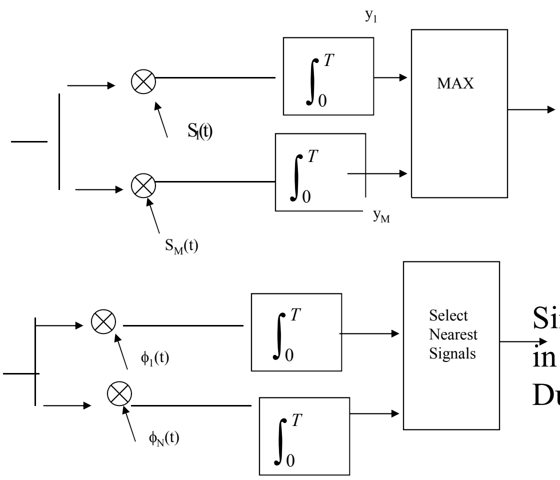
(M - 1) possible symbol errors. $M/2$ bit errors for each bit position

$\frac{M/2}{M-1}$ = conditional prob. of bit error given a symbol error

$P_e = \frac{M/2}{M-1} P(\varepsilon)$ valid if the distance between any two symbols is equal

( e.g. orthogonal signals)

# Coherent detection of M-ary signals




Simpler
in practice
Due to N<=M

43

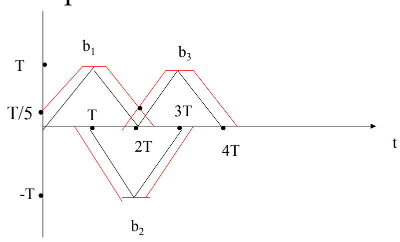
# Intersymbol interference

Causes:

- Band-limited channel
- Frequency selective fading (delay spread>symbol interval)- see next notes

44

## Example




Sample at t = 2T

$$b_2 T + b_1 \cdot \frac{T}{5} + b_3 \cdot \frac{T}{5}$$

$$signal \quad -T \pm \frac{T}{5} \pm \frac{T}{5} = \frac{-3T}{5}, -T, -T, \frac{-7T}{5}$$

$$\overline{P_e} = \frac{1}{4}[Q(\frac{3T/5}{\sigma}) + 2Q(\frac{T}{\sigma}) + Q(\frac{7T/5}{\sigma})] \quad \text{(p.26)}$$

$$= E\{Q(\frac{|signal|}{\sigma})\} \neq Q(\frac{T}{\sigma}) = Q(\frac{E\{|signal|\}}{\sigma})$$

45

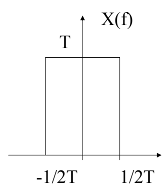
Condition for No ISI

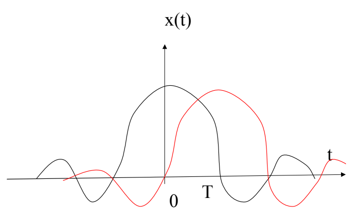
$x(t)$ : pulse shape after filtering in Rx

$$x(t) = \sum_{n=-\infty}^{\infty} x(nT) \frac{\sin \pi(t-nT)/T}{\pi(t-nT)/T}$$

to avoid ISI $\Rightarrow$ $x(nT) = 0 \;\; \forall\, n \neq 0$

Example $\quad X(f) = \begin{cases} T & |f| < \frac{1}{2T} \\ 0 & \text{otherwise} \end{cases}$




46

# Nyquist Criterion for No ISI

problem : if sampling time is not correct $\sum \left|\frac{\sin(\frac{\pi t}{T})}{(\frac{\pi t}{T})}\right| \approx \sum \frac{1}{(\frac{\pi t}{T})}$ diverge

Let $\frac{1}{T} < 2W$ ( X(f) is bandlimited to [-W, W] )

Let $Xeq(f) = \sum_{n=-N}^{N} X(f+\frac{n}{T})$ $|f| < \frac{1}{2T}$ aliasing DTFT

If $x(kT) = \int_{\frac{-1}{2T}}^{\frac{1}{2T}} X_{eq}(f)\cdot e^{j2\pi fkT}\,df = 0$ for $k \neq 0$

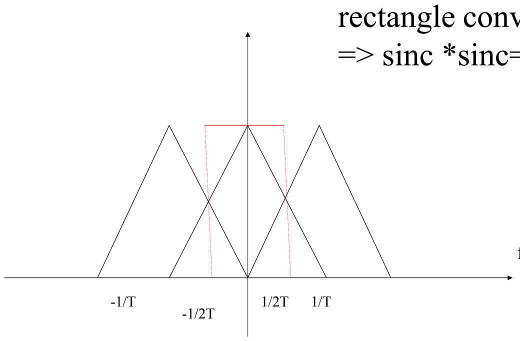
then

$X_{eq}(f) = \begin{cases} constant & |f| \le \frac{1}{2T} \\ 0 & otherwise \end{cases}$ (Nyquist criterion for No ISI)

47

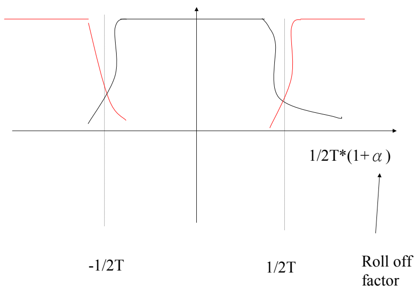
# Example 1: sinc$^2$

rectangle convol rectangle in freq domain
=> sinc *sinc=sinc$^2$ in time domain




48

# Example 2: raised cosine(RC)




Commonly used
Ex. Vector signal analyzer
Can set 1)TX RC 2) RX RC
3) TX sqrt(RC) RC sqrt(RC)

49

# Linear Algebra Review

國中解聯立方程式的延伸

1) Pseudo inverse- NWTT5 MIMO spatial multiplexing
2) QR decomposition
3) Singular Value Decomposition (SVD)- optimal MIMO precoding for rank deficiency NWTT8 p.31 (4G TMs)

50

# Review of linear algebra

- $\vec{X} = (X_1, X_{2,} \cdots, X_n) \in \Re^n$
  $\vec{Z} = (Z_1, Z_{2,} \cdots, Z_n) \in C^n$ .
- N vectors are linearly independent satisfying following :
  $$\sum_{i=1}^{n} a_i \vec{X}_i = 0 \quad iff \quad a_i = 0 \quad \forall \quad i$$
- Span of a set of vectors
  $$\{\vec{S} \mid \vec{S} = \sum_{i=1}^{n} a_i \vec{X}_i, \forall a_i \in R\}$$

51

- A basis for a vector space is a set of vectors with properties :
  * It is linear independent. (not nec orthogonal)
  * It spans the space.
- Orthonormal basis
  $\vec{X}_1, \vec{X}_2, \cdots, \vec{X}_n$ is orthogonal and length 1.
- $n \times m$ matrix : a mapping $R^m \to R^n$
  $$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ X_m \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

52

domain of the matrix M = $R^m$

range of the matrix M = $\{\vec{y} \mid M\vec{X} = \vec{y}, \forall \vec{X} \in R^m\}$

- Null space of M = $\{\vec{X} \mid M\vec{X} = 0\}$
  nullity = dimension of nullspace.
- Rank = dimension of the range space.
  Rank + nullity = dimension of domain.
- An nxn matrix M is invertible iff it has full rank : $\det M \neq 0,\ eigenvalues \neq 0.$

53

- Ex.

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 2 & 3 & 2 \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} X_1 + X_3 = 0 \\ X_2 = 0 \end{cases}$$

$$\Rightarrow \begin{bmatrix} X_1 \\ 0 \\ -X_1 \end{bmatrix} \Rightarrow \vec{X} = X_1\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

Only one variable, nullity = 1.
Rank = 2.

54

- Inner product- operation on 2 vectors with the following properties:

$$\vec{X}\cdot\vec{X} \ge 0, equality \quad holds \quad iff \quad \vec{X} = 0$$

$$\vec{X}\cdot\vec{Y} = \vec{Y}\cdot\vec{X} (X, Y \in C \Rightarrow add \quad complement: \ \vec{X}\cdot\vec{Y}^*)$$

$$(\vec{X}+\vec{Y})\cdot\vec{Z} = \vec{X}\cdot\vec{Z} + \vec{Y}\cdot\vec{Z}$$

$$(r\vec{X})\cdot\vec{Y} = r(\vec{X}\cdot\vec{Y})$$

55

- Ex.
  $\vec{X}\cdot\vec{Y} = \vec{X}^T \cdot M \cdot \vec{Y}$ is an inner product if M is positive definite ($\vec{X}^T \cdot M \cdot \vec{X} \ge 0$ , equality holds $iff \quad \vec{X} = 0$)
- Ex.
  $\vec{X}\cdot\vec{Y} = \sum_{i=0}^{n} X_i Y_i$
- Ex.
  $\vec{X}\cdot\vec{Y} = \int_0^T X(t)Y(t)dt$

56

- Norm ||X||- vector operation with the following properties

$$\| \vec{X} \| \geq 0, \textit{equality holds iff } \vec{X} = 0.$$

$$\| r\vec{X} \| = r \| \vec{X} \|.$$

$$\| \vec{X} + \vec{Y} \| \leq \| \vec{X} \| + | \vec{Y} \|.$$

Ex. $\| \vec{X} \| = \sqrt{\vec{X} \cdot \vec{X}}$

Ex. $\sqrt[4]{\sum_{i=1}^{n} | X_i |^4}$

57

# Approaches to solve Mx≈b

- Normal equations-quick and dirty: pseudo inverse, numerically unstable
- QR- standard in libraries uses orthogonal decomposition
- SVD - decomposition which also gives indication how linear independent columns are

Quick and Dirty Approach

Multiply by $M^T$ to get the **normal equations**:

$$M^T M x = M^T b$$

However, sometimes $M^T M$ can be *nearly singular or singular.*

Consider the matrix $M = \begin{bmatrix} 1 & 1 \\ e & 0 \\ 0 & e \end{bmatrix}$

The matrix $M^T M = \begin{vmatrix} 1+e^2 & 1 \\ 1 & 1+e^2 \end{vmatrix}$

becomes singular if e is less than the square root of the machine precision.

# Goal of QR: a more stable approach

$$\underset{m\times n}{M} = \underset{m\times m}{Q}\underbrace{\begin{bmatrix} R \\ O \end{bmatrix}}_{m\times n} = Q\begin{bmatrix} ? & ? & \cdots & ? \\ 0 & \ddots & & \vdots \\ \vdots & 0 & \ddots & \vdots \\ \vdots & & 0 & ? \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{array}{l} \left.\vphantom{\begin{matrix} ? \\ 0 \\ \vdots \\ \vdots \end{matrix}}\right\} \text{R: nxn, upper tri.} \\ \left.\vphantom{\begin{matrix} 0 \\ 0 \end{matrix}}\right\} \text{O: (m-n)xn, all zeros} \end{array}$$

Ex. 3 (n) unknows 5 (m) equations because of noise

60

# Reformulating Least Squares using QR

$\|r\|^2 = \|b - Mx\|^2 = \left\|b - Q\begin{bmatrix} R \\ O \end{bmatrix}x\right\|^2 = \left\|Q^T b - Q^T Q\begin{bmatrix} R \\ O \end{bmatrix}x\right\|^2$

because Q is orthonormal ($Q^T Q = I$) $\left\|Q \quad x\right\|^2 = \|x\|^2$ (see next page)

$= \left\|\begin{matrix} c_1 - Rx \\ c_2 \end{matrix}\right\|^2$ because $Q^T b = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$

$= \|c_1 - Rx\|^2 + \|c_2\|^2$

$>= \|c_2\|^2$ minimum error if we choose $Rx = c_1$

**Numerically stable because R is upper triangular (even if M has no inverse)**

**Can use** 國中數學 解聯立方程式 消去法

2x+3y=5

x+y =2

=>2x+3y=5

y=1

61

QR for least square

A matrix Q is **orthonormal** if $Q^T Q = I$.

If Q is orthogonal then for any x, $\|Qx\|^2 = \|x\|^2$,

that is an orthogonal matrix preserves the 2 norm.

Examples of orthogonormal matrices:

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} \cos y & \sin y \\ -\sin y & \cos y \end{pmatrix}$ Givens rotations

for some angle y

Most good least squares solvers use the QR approach.

***In Matlab: x= M \b.***

## Singular Value Decomposition

The singular value decomposition(SVD) of a matrix A is given by

M= U **Σ** V$^T$

mxn=mxm mxn nxn

where U and V are unitary matrices (座標轉換 e.g. 圓座標 <=>直角座標)and S (座標值) is a rectangular diagonal matrix.

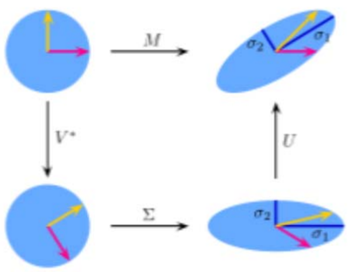
The columns of U are left singular vectors (orthonormal basis of $K^m$, $K$=$R$ real or $C$ complex)

The columns of V are right singular vectors (orthonormal basis of $K^n$)

The diagonal entries of **Σ** are singular values of M.

# Visualization of the SVD of a 2D matrix M

- The SVD decomposes **M** into three simple transformations: an initial rotation V*, a scaling **Σ** along the coordinate axes, and a final rotation U. The lengths $\sigma_1$ and $\sigma_2$ of the semi-axes of the ellipse (橢圓) are the singular values of M.



$M = U \cdot \Sigma \cdot V^*$

- The left-singular vectors of **M** are a set of orthonormal eigenvectors of **MM***.
- The right-singular vectors of **M** are a set of orthonormal eigenvectors of **M*M**.
- The non-zero singular values $\sigma_i$ of **M** (found on the diagonal entries of **Σ**) are the square roots of the non-zero eigenvalues of both **M*M** and **MM***.

65

- The linear transformation

$$\begin{cases} T : K^n \to K^m \\ x \mapsto \mathbf{M}x \end{cases}$$

- has a particularly simple description with respect to these orthonormal bases: we have

$$T(\mathbf{V}_i) = \sigma_i \mathbf{U}_i, \qquad i = 1, \ldots, \min(m, n),$$

66

where $\sigma i$ is the $i$-th diagonal entry of $\mathbf{\Sigma}$, and $T(\mathbf{V}i) = 0$ for $i > \min(m,n)$.

- The geometric content of the SVD theorem can thus be summarized as follows: for every linear map $T : K^n \rightarrow K^m$ one can find orthonormal bases of $K^n$ and $K^m$ such that $T$ maps the $i$-th basis vector of $K^n$ to a non-negative multiple of the $i$-th basis vector of $K^m$, and sends the left-over (|m-n|) basis vectors to zero. With respect to these bases, the map $T$ is therefore represented by a diagonal matrix with non-negative real diagonal entries.

67

## SVD for least squares

If M is an m x n matrix of rank n, then

$$M = U \mathbf{\Sigma} V^T = [\, U_1 \quad U_2 \,] \begin{bmatrix} \Sigma_1 \\ 0 \end{bmatrix} V^T = U_1 \mathbf{\Sigma}_1 V^T$$

mxm mxn $\Sigma$

where $U_1$ has the first n columns of U and $\mathbf{\Sigma}_1$ is n x n.

The solution to the least squares problem $Ax \approx b$ is given by

$x = V\Sigma_1^{-1}U_1^Tb$

Requires 4 to 10 times more work than QR but shows dependencies in model.

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{U} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} 2 & 0 & 0 & 0 & \mathit{0} \\ 0 & 3 & 0 & 0 & \mathit{0} \\ 0 & 0 & \sqrt{5} & 0 & \mathit{0} \\ 0 & 0 & 0 & \mathbf{0} & \mathit{0} \end{bmatrix}$$

$$\mathbf{V}^* = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \sqrt{0.2} & 0 & 0 & 0 & \sqrt{0.8} \\ 0 & 0 & 0 & 1 & 0 \\ -\sqrt{0.8} & 0 & 0 & 0 & \sqrt{0.2} \end{bmatrix}$$

# Example

- Notice $\Sigma$ is zero outside of the diagonal (grey italics) and one diagonal element is zero (red bold)

# Application

- [Ali17] MIMO-NOMA:
- Choose the left singular vector corresponding to largest singular value to represent the MIMO channel vectors of multiple users in a NOMA cluster,
- Max replaces sum (reduced dimension)

# ML/MAP Estimation and Detection

Example: x= $\theta$ +n, $\theta$ =1 or -1, n=AWGN

MAP and ML



## Introduction

- **Estimation theory plays a major role in the design of statistical signal processing systems. Estimation problems are often delineated into deterministic (or classical) parameter estimation and random (or Bayesian) parameter estimation.**
- **1) classical estimation, observations are random and parameters are regarded as unknown constant (deterministic) values.**
- **2) Bayesian parameter estimation, the parameters are also viewed as random, and our prior knowledge of their behavior is expressed through an a priori density.**
  - **Maximum A Posteriori (MAP) – Bayesian (with prior)**
  - **Maximum Likelihood (ML) -- Classical**

**Maximum Likelihood => minimum Euclidean distance in AWGN (demodulation, Viterbi decoding etc.)**

Assume AWGN channel with vari nace $\sigma^2$

(y = x + AWGN)

log likelihood function

$$= \log \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(y-x)^2}{2\sigma^2}}$$

$$= \log \frac{1}{\sqrt{2\pi\sigma^2}} - \frac{1}{2\sigma^2}(y-x)^2$$

bit metric $= (y-x)^2$

(Euclidean 歐幾里德 distance) $^2$ !

73

Maximum a posterior probability (MAP) including prior probability (used in turbo decoding, soft-input soft-output turbo principle etc.)

$$p(x=+1\,|\,y) = p(y\,|\,x=+1)p(x=+1)/p(y)$$

$$p(x=-1\,|\,y) = p(y\,|\,x=-1)p(x=-1)/p(y)$$

a posterior probability (APP)

= likelihood function * prior probability

If prior probability = constant (1/2)

then maximum a posterior probability = maximum likelihood

74