BIOE 504

→ Measurement System Modeling

$$\underbrace{\underset{\substack{\uparrow \\ f(t')}}{\text{Object}} \rightarrow \underset{\mathcal{H}}{\text{Measurement System}} \rightarrow \underset{g(t)}{\text{Data}}}_{\text{Acquisition Stage}} \rightarrow \text{Display Recovery} \rightarrow \text{Object}$$

$g(t) = \int_{-\infty}^{+\infty} h(t, t') f(t')\, dt'$ (linear) LTI / LSI

$f(t') \rightarrow \boxed{\mathcal{H}} \rightarrow g(t)$

$f(t'-t_0) \rightarrow \boxed{\mathcal{H}} \rightarrow g(t-t_0)$

Define $h(t')$ Such that $g(t) = \underset{\text{conv}}{[h \overset{\Delta}{*} f]}(t)$

Examples:

$g(t) = \int_{-\infty}^{t} f(t')\, dt'$ (System #1) $f_3(t') = \alpha f_1(t') + \beta f_2(t')$

$g_3(t) = \int_{-\infty}^{t} f_3(t')\, dt' = \alpha g_1(t) + \beta g_2(t)$ ⟶ linear

$g_1'(t) = \int_{-\infty}^{t} f_1(t'-t_0)\, dt' \overset{w = t'-t_0}{=\!=} \int_{-\infty}^{t-t_0} f_1(w)\, dw = \int_{-\infty}^{t-t_0} f(t')\, dt' = g_1(t-t_0)$

System #2 $g(t') = f^2(t')$ Definition → No linear

System #3 $g(t') = f(t') + \underset{0}{a(t')}$ $\alpha f_1(t') + \beta f_2(t') \Rightarrow$ Creat $2a(t')$

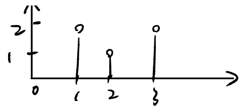
Sampling

$g(t) = \int_{-\infty}^{+\infty} h(t-t') f(t')\, dt'$ $g[m] = g(mT) = \int_{-\infty}^{+\infty} h(mT-t') f(t')\, dt'$

Discretize $f(t')$ as $f[n] = f[nT']$ $n = -\infty, \cdots, +\infty$

$g[m] = \sum_{n=-\infty}^{+\infty} h(mT - nT') f[nT'] = \sum_{n=-\infty}^{+\infty} h[m-n] f[n]$

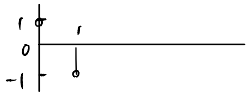
$\Rightarrow f[n]$



$[1,2,1,2]^T$

$$g[m] = \sum_{n=-\infty}^{+\infty} h[m-n]\, f[n]$$

$h[n]$



$[1,-1]^T$

$$g[0] = \sum_{n=-\infty}^{+\infty} h[-n] f[n] = h(-1)f(1) + h(0)f(0) + h(1)f(-1) = 1$$

$\vdots \quad g[1]=1 \quad g[2]=-1 \quad g[3]=1 \qquad H_{i+1,j+1} = H_{ij}$

$$\vec{f} \triangleq \begin{bmatrix} f[0] \\ f[1] \\ f[2] \\ f[3] \end{bmatrix} \qquad \vec{g} = \begin{bmatrix} g[0] \\ g[1] \\ g[2] \\ \vdots \\ g[4] \end{bmatrix} \qquad H \triangleq \underbrace{\overbrace{\begin{bmatrix} h(0) & 0 & 0 & 0 \\ h[1] & h[0] & 0 & 0 \\ h[2] & h[1] & h[0] & 0 \\ 0 & h[2] & h[1] & h[0] \\ 0 & 0 & h[2] & h[1] \end{bmatrix}}^{\text{Toeplitz}}}_{5\times 4} \underbrace{\begin{bmatrix} f[0] \\ f[1] \\ f[2] \\ f[3] \end{bmatrix}}_{4\times 1}$$

$$\vec{f} \in R^{\overset{N}{4}}, \quad \vec{h} \in R^{\overset{M}{2}} \qquad \vec{g} \in R^{\overset{N+M-1}{4+2-1}} = R^5 \qquad H \in R^{(N+M-1)\times N}$$

Any linear shift invariant System can be represent by a Toeplitz Matrix. Formed from impulse response function

1D convolution (discrete) 2D convolution

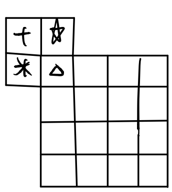
$$g[m] = \sum_{n=-\infty}^{+\infty} h[m-n]\, f[n] \qquad g[m,n] = \sum_{m'=-\infty}^{+\infty} \sum_{n'=-\infty}^{+\infty} h[m-m', n-n']\, f[m', n']$$

D-Dimensional Convolution

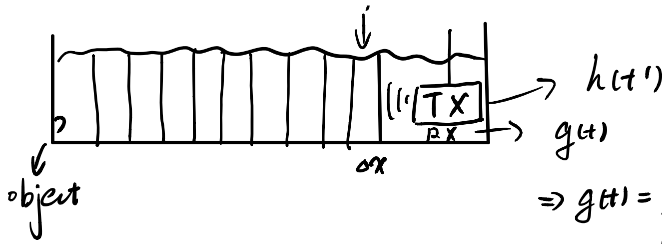
$$g[n_1, n_2, \cdots n_D] = \sum_{m_1, m_2 \cdots m_D} h[n_1-m_1, n_2-m_2, \cdots n_D-m_D]\, f[m_1, m_2, \cdots m_D]$$

$$H = \underbrace{\left[\begin{array}{cccccccccc} \triangle & 0 & 0 & 0 & & \cdots & & & & 0 \\ \star & \triangle & 0 & 0 & & \cdots & & & & 0 \\ 0 & \star & \triangle & 0 & & \cdots & & & & 0 \\ 0 & 0 & \star & \triangle & & \cdots & & & & 0 \\ \ast & \cdots & & & \triangle & \cdots & & & & \\ + & \ast & \cdots & & \star & \triangle & \cdots & & & \\ 0 & + & \ast & 0 & 0 & \star & \triangle & & & \\ 0 & 0 & + & \ast & 0 & 0 & \star & \triangle & & \end{array}\right]}_{25\times 16} \quad (\text{Toeplitz}) \qquad \vec{f} = \underbrace{\begin{bmatrix} f[0,0] \\ f[1,0] \\ f[2,0] \\ f[3,0] \\ f[0,1] \\ f[1,1] \\ f[2,1] \\ f[3,1] \\ \vdots \end{bmatrix}}_{16\times 1} \begin{matrix} \text{Column 1} \\ \\ \text{Column 2.} \end{matrix}$$

Block Toeplitz Matrix



$f(n\Delta x)\ (n=0,\cdots,N)$

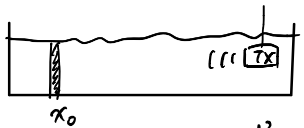


$f(t') \underbrace{f(n\Delta x)}_{\text{density}} \underbrace{\Delta x}_{\text{width of piece}}$

$g(t) = \sum_{n=0}^{N-1} h(t-t') f(n\Delta x)\Delta x \quad t' = \frac{2n\Delta x}{v}$

$\Rightarrow g(t) = \sum_{n=0}^{N-1} h(t - \frac{2n\Delta x}{v}) f(n\Delta x)\Delta x$

$g(t) = \int_{-\infty}^{+\infty} h(t - \frac{2x}{v}) f(x)\, dx$

① 

$f(x) = \delta(x) + R\,\delta(x - x_0)$
$h(t) = Ae^{-\frac{t^2}{2\sigma^2}} \sin(\Omega_0 t)$

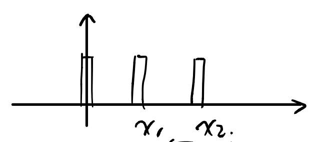
$g(t) = \int_{-\infty}^{+\infty} Ae^{-\frac{(t-\frac{2x}{v})}{2\sigma^2}} \sin[\Omega_0(t - \frac{2x}{v})][\delta(x) + R\delta(x - x_0)]\, dx$

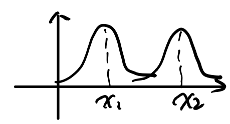
$g(t) = Ae^{-\frac{t^2}{2\sigma^2}} \sin(\Omega_0 t) + RAe^{\frac{-(t-\frac{2x_0}{v})^2}{2\sigma^2}} \sin(\Omega_0(t - \frac{2x_0}{v}))$

$\downarrow$ demodulation

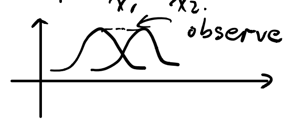
$g(t) = Ae^{-\frac{t^2}{2\sigma^2}} + ARe^{-\frac{(t-\frac{2x_0}{v})^2}{2\sigma^2}}$

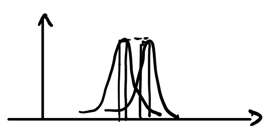
② $f(x) = R_1\delta(x - x_1) + R_2\delta(x - x_2)$
$g(t) = R_1 h(t - \frac{2x_1}{v}) + R_2 h(t - \frac{2x_2}{v})$





$h(t) = Ae^{-\frac{t^2}{2\sigma^2}}$ if $\sigma\uparrow$ too large $\Rightarrow$





$\Delta x = x_1 - x_2$

$\Delta x \gg \sigma$ better chance to distinguish 2 signals (resolution)

$f(x) = \sum_{i=0}^{N-1} R_i b_i(x)$ e.g., $b_i(x) = \delta(x - i\Delta x)$

$g(t) = \int_{-\infty}^{+\infty} h(t - \frac{2x}{v}) \sum_{i=0}^{N-1} R_i b_i(x)\, dx = \sum_{i=0}^{N-1} R_i \int_{-\infty}^{+\infty} h(t - \frac{2x}{v}) b_i(x)\, dx$

if estimate $\{\hat{R}_i\}_{i=1}^{N-1}$, then $\hat{f}(x)$ can be reconstruct

$\hat{f}(x) = \sum_{i=0}^{N-1} \underbrace{\hat{R}_i}_{\text{from data}} \underbrace{b_i(x)}_{\text{known}}$

collect data at $mT \leftarrow$ sampling interval
$g[m] = g[mT]$

$g[mT] = \sum_{i=0}^{N-1} R_i \underbrace{\int_{-\infty}^{+\infty} h(mT - \frac{2x}{v}) b_i(x)\, dx}_{a_{mi}}$

$g[m] = \sum_{i=0}^{N-1} a_{mi} R_i$

Define a vector $\vec{g} = \begin{bmatrix} g[0] \\ g[1] \\ \vdots \\ g[M-1] \end{bmatrix}$ $\vec{R} = \begin{bmatrix} R_0 \\ R_1 \\ \vdots \\ R_{N-1} \end{bmatrix}$

$\vec{g}_{M\times 1} = A \vec{r}_{N\times 1}$ $(M \gg N)$ $\Rightarrow A_{mi} = a_{mi}$

$\vec{g}_m = \sum_{i=0}^{N-1} a_{mi} \vec{r}_i$

if we can get $\hat{r} = A^{+} \vec{g}$, then $\hat{f}(x) = \sum_{i=0}^{N-1} \hat{R}_i b_i(x)$

some kind of inversion. (M, N relationship

key point: estimate A

let A has an inversion matrix)

"Decomposition"

Example: $f(x) = \sum_{i=1}^{N-1} c_i b_i(x)$ basis element

- Independence

$\vec{v} = [v_1, v_2, \cdots v_N]^T$ collection of measurements of some kind of object

$\vec{v} \in R^N, \phi^N$

Given $\vec{v}_1, \vec{v}_2 \in R^N$ $\vec{v}_1, \vec{v}_2$ are independent when there is no such a scalar $c \neq 0$

$c\vec{v}_1 + \vec{v}_2 = 0$

In other word, if $\exists c \neq 0$, s.t. $c\vec{v}_1 + \vec{v}_2 = 0$, then not independent

Generalize to any M vectors $\{\vec{v}_i\}_{i=1}^{M}$ $\vec{v}_i \in R^N$

If there isn't a set of nonzero coefficients $\{c_i\}_{i=1}^{M}$

s.t. $\sum_{i=1}^{M} c_i \vec{v}_i = 0$

then $\{\vec{v}_i\}_{i=1}^{M}$ are linear dependent

assume $\{\vec{v}_i\}$ are linear dependent

then $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + C_k \vec{v}_k + \cdots = 0$

let $C_k \neq 0$, $\vec{v}_k = \frac{c_1}{c_k}\vec{v}_1 + \frac{c_2}{c_k}\vec{v}_2 + \cdots + \frac{c_{k-1}}{c_k}\vec{v}_{k-1} + \frac{c_{k+1}}{c_k}\vec{v}_{k+1} + \cdots$

→ Orthogonality

Given $\vec{v}_1, \vec{v}_2 \in R^N$, they are orthogonality if $\langle \vec{v}_1, \vec{v}_2 \rangle = 0$

For a set of vectors $\{v_i\}_{i=1}^{M}$

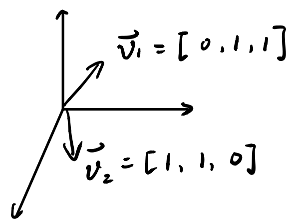
if $\forall \vec{v}_1, \vec{v}_2 \in \{v_i\}_{i=1}^{M}$ are orthogonal, $\{v_i\}_{i=1}^{M}$ is a set of orthogonal vectors

→ Uncorrelated vectors

$\vec{v}_1, \vec{v}_2 \in R^N$ $\bar{v}_1 = \frac{\sum_{i=1}^{N} v_{1,i}}{N}$ $\bar{v}_2 = \frac{\sum_{i=1}^{N} v_{2,i}}{N}$

$\vec{v}_1, \vec{v}_2$ are uncorrelated

if $<\vec{v}_1 - \bar{v}_1 \vec{1}, \vec{v}_2 - \bar{v}_2 \vec{1}> = 0$



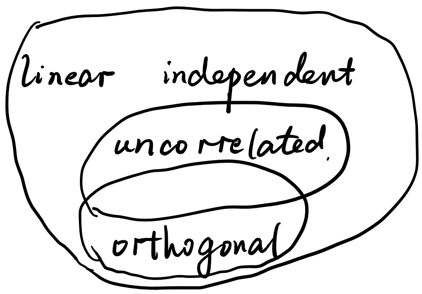
linear independent

not orthogonal $<\vec{v}_1, \vec{v}_2> = 1$

$\vec{v}_1 - \frac{2}{3}\vec{1} = [-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}]$

$\vec{v}_2 - \frac{2}{3}\vec{1} = [\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}]$

$\Big\}$ corretaled



$\rightarrow$ Random variable

① X (stochatic object) X can take values from a set of values

each x has probability associated with it.

$P(X = x_i) = P_i \quad (i = 1, \cdots, N)$ (Probability measure)

② P_i probability distribution

If $x \in R$ (possible value, continously distributed along a line)

$p_i(x)$: probability distribution function (PDF)

For any continuous functions $\{f_i(x)\}_{i=1}^{M}$ $f_i(x) \in R$

$\sum_{i=1}^{M} \alpha_i f_i(x) = 0 \Leftrightarrow$ all α_i is zero, then $\{f_i(x)\}$ is linear indipendent

Any two functions $f_i(x)$ $f_j(x)$ $<f_i(x), f_j(x)> = \int_{-\infty}^{+\infty} f_i(x) f_j(x)\, dx = 0$

then $f_i(x)$ & $f_j(x)$ are othrogonal.

$\int_{-\infty}^{+\infty} (f_i(x) - \bar{f}_i)(f_j(x) - \bar{f}_j)\, dx = 0$

Ramdom Variables

The value of X can be from

① $S = \{x_i\}_{i=1}^{N}$: X can take anyone of the N possible values
② $S = R$ (real line) : X can be real numbers.

For ① $p(X=x_i) = p_i$, $p_i \in [0,1]$, $\sum p_i = 1$ (PMF)

② $p(X \le x_i) = CDF(X) = F(x_i)$ $F'(x) = f(x_i) \triangleq PDF$ ← density

PDF: $\int_{-\infty}^{+\infty} f(x_i)\, dx = 1$

Expectation of X : $E[X] \triangleq \begin{cases} \int_{-\infty}^{+\infty} x f(x)\, dx \\ \sum x_i p_i(x) \quad \text{discretize} \end{cases}$

Variance of $X = E[X^2] - E^2[X]$ $E[X^2] = \int x^2 f(x)\, dx$ 2^{nd} mom -

Covariance of RV X & Y

$E[X - E(X)][Y - E(Y)] = cov(X, Y)$

→ Independence X & Y are a R.V.

Joint PDF / distribution $P(X \le x, Y \le y) = CDF(X, Y)$

$CDF' = p(x, y)$ Joint PDF = Derivative of Joint CDF

Independent.

$P(X \le x, Y \le y) = P(X \le x)\, P(Y \le y)$

For Joint PDF $f(x,y) = f(x) f(y)$

$E[XY] \triangleq \int_{-\infty}^{+\infty} xy f(x,y)\, dx dy$ ($\langle X, Y \rangle$)

if $E[X,Y] = 0$, Then X & Y are orthogonal.

If $cov(X,Y) = 0$, then X & Y are uncorrelated

→ Signal decomposition (Data representation)

⇒ vector space & subspace

$V = \{\vec{v}_i, \vec{v}_i \in R^n\}$ a set of vectors

Any $\vec{v}_1, \vec{v}_2 \in V$, $\alpha \vec{v}_1 + \beta \vec{v}_2 \in V$

⇒ subspace $S \subseteq V$ (a subset of V) such that Any $\vec{v}_1$ & $\vec{v}_2 \in S$ $\alpha\vec{v}_1 + \beta\vec{v}_2 \in S$

$\Rightarrow$ Basis: for a subspace S

If I have a set of vectors $\{\vec{b}_i\}_{i=1}^{D}$, such that

① For any $\vec{v} \in S$ $\vec{v} = \sum_{i=1}^{D} \alpha_i \vec{b}_i$

② $\{\vec{b}_i\}_{i=1}^{D}$ are linear independent

if ① & ② hold, then $\{b_i\}_{i=1}^{D}$ is a basis of the subspace S

☆ D is the dimension of the subspace

e.g. $\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in R^3$ $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = a\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

Function $f(x) \in R$, continuous

$f(x) = a + bx + cx^2$

$\{b_1(x)\ b_2(x)\ b_3(x)\}$ For any $f(x) = a + bx + x^2 \in \{2^{nd}$ order polynomials$\}$

$f(x) = \sum_{i=1}^{3} \alpha_i b_i(x)$

$\rightarrow$ For any general function $f(x)$ or vector $\vec{v}$

express $f(x)$ or $\vec{v}$ as follow:

① $f(x) = \sum_{i=1}^{N \text{ or } \infty} \alpha_i b_i(x)$ ② $\vec{v} = \sum_{i=1}^{D} \alpha_i \vec{b}_i$ $\vec{b}_i \in R^N$, $\vec{v} \in R^N$

or "$\approx$" approximate

$\Rightarrow$ orthonormal basis orthogonal + $\|\vec{b}_i\|_2^2 = 1$ (vector) or $\int_{-\infty}^{+\infty} b_i^2(x)dx = 1$ (energy)

$\{\vec{b}_i\}_{i=1}^{D}$ or $\{b_i(x)\}$

$\alpha_j = \langle f(x), b_j(x) \rangle$

$= \langle \sum_{i=1}^{D} \alpha_i b_i(x), b_j(x) \rangle = \sum_{i=1}^{D} \langle \alpha_i b_i(x), b_j(x) \rangle = \alpha_j \langle b_j(x), b_j(x) \rangle = \alpha_j$

$\alpha_k = \int f^*(t) b_k(t)\, dt$ coefficient are simply inner product of then orignal function and individual basis function

→ Peoriodic function ($f(t)$ or $f(x)$)

Fourier series $b_k(t) = e^{i2\pi k u_0 t}$ u_0: fundemental fequency

ku_0: harmonic of then (fundemental frequency)

integral $(-\infty, +\infty)$

$$f(t) = \sum_{-\infty}^{+\infty} \alpha_k e^{i2\pi k u_0 t}$$ (Fourier series representation)

$$e^{i2\pi k u_0 t} = \cos(2\pi k u_0 t) + i\sin(2\pi k u_0 t)$$

$$\langle b_\ell(t), b_k(t) \rangle = \int_{-\infty}^{+\infty} e^{-i2\pi \ell u_0 t} \cdot e^{i2\pi k u_0 t} = \frac{1}{T_0}\int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} e^{i2\pi(k-\ell)u_0 t}\, dt$$

$$* = \frac{1}{T_0}\int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} \{\cos[2\pi(k-\ell)u_0 t] + \underbrace{i\sin[\cancel{2\pi(k-\ell)u_0 t}]}_{0}$$

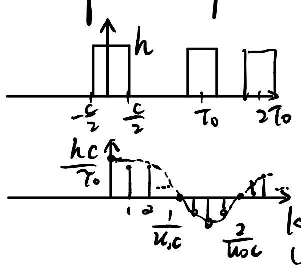
case ① $k \neq \ell$ $* = 0$ $[-\frac{T_0}{2}, \frac{T_0}{2}]$ cover multiple perıods of $\cos[2\pi(k-\ell)u_0 t]$

case ② $k = \ell$ $* = 1$

$$\alpha_k = \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} f(t)\, e^{-i2\pi k u_0 t}\, dt$$ (FS coefficient Forward FS)

$$f(x) = \sum_k \alpha_k e^{i2\pi k u_0 x}$$

Example : $f(t) = h\,\mathrm{rect}\left(\frac{t}{c}\right)$ $-\frac{T_0}{2} \le t \le \frac{T_0}{2}$ (periodic)



$$f(t) = \sum \alpha_k e^{i2\pi k u_0 t} \quad u_0 = \frac{1}{T_0}$$

$$\alpha_k = \frac{1}{T_0}\int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} f(t)\, e^{-i2\pi k u_0 t}$$

$$= \frac{1}{T_0}\int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} h e^{-i2\pi k u_0 t} = \frac{h}{T_0}\int_{-\frac{c}{2}}^{\frac{c}{2}} [\cos(2\pi k u_0 t)] - \underbrace{i\sin(\cancel{2\pi k u_0 t})}_{0}]\, dt$$

$$= \frac{h}{T_0}\frac{1}{2\pi k u_0}\sin(2\pi k u_0 t)\Big|_{-\frac{c}{2}}^{\frac{c}{2}} = \frac{hc}{T_0}\,\mathrm{sinc}(k u_0 c)$$

(frequency domain)

{ ① magnitude of α_k decaying $f(t) \approx \sum_{k=1}^{N} \alpha_k e^{i2\pi k u_0 t}$

② multiple zeros-crossing

→ Fourier Transform (FT) (CTFT) Motivation: $f(t)$ non-periodic

FS: $$\alpha_k = \frac{1}{T_0}\int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} f(t)\, e^{-i2\pi k u_0 t}\, dt$$

$\Downarrow$ $T_0 \to \infty$ $\frac{1}{T_0} \to du$ $ku_0 \to u$

$$F(u) = \int_{-\infty}^{+\infty} f(t)\, e^{-i2\pi u t}\, dt. \qquad f(t) = \int_{-\infty}^{+\infty} F(u)\, e^{i2\pi u t}\, du.$$

(Forward FT) (Inverse FT)

$$\mathcal{F}[f(t)] = F(u) \; ; \; f(t) = \mathcal{F}^{-1}[F(u)]$$

$\Rightarrow$ for any function $f(t)$. s.t.

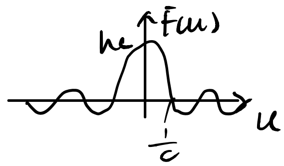
$$\int_{-\infty}^{+\infty} |f(t)|\,dt < \infty$$

Example: $f(t) = h\,\mathrm{rect}\left(\frac{t-b}{c}\right)$

$$F(u) = \int_{-\infty}^{+\infty} f(t)\, e^{-i2\pi ut}\,dt = \int_{-\frac{c}{2}}^{\frac{c}{2}} h e^{-i2\pi u(t'+b)}\,dt'$$

$$= h e^{-i2\pi ub} \int_{-\frac{c}{2}}^{\frac{c}{2}} e^{-i2\pi ut'}\,dt' = hc\,e^{-i2\pi ub}\,\mathrm{sinc}(uc)$$

Linear phase



$F(u)$. Fourier spectrum of $f(t)$

$|F(u)|$: magnitude spectrum.

$\angle F(u)$: Phase spectrum.

$\rightarrow$ Properties of FT

① linearity $\mathcal{F}[\alpha f_1(t) + \beta f_2(t)] = \alpha F_1(u) + \beta F_2(u)$

② FT of delta function $\mathcal{F}[\delta(t)] = \int_{-\infty}^{+\infty} \delta(t)\, e^{-i2\pi ut}\,dt = 1$

$\mathcal{F}[c\,\delta(t)] = c$

$\mathcal{F}[c] = c\,\delta(t)$

$\mathcal{F}[\delta(t-t_0)] = \int_{-\infty}^{+\infty} \delta(t-t_0)\, e^{-i2\pi ut}\,dt$

$= e^{-i2\pi ut_0}$

③ Convolution Property

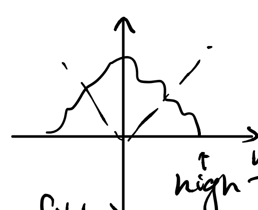
$\mathcal{F}[f_1(t) * f_2(t)] = F_1(u)\cdot F_2(u)$ $\mathcal{F}[f_1(t) f_2(t)] = F_1(u) * F_2(u)$

$\rightarrow$ shifting property

$\mathcal{F}[f(t-t_0)] = F(u)\, e^{-i2\pi ut_0}$

$\rightarrow$ Differentiation

$\mathcal{F}\left[\frac{df(t)}{dt}\right] = j2\pi u F(u) \xrightarrow{\text{magnitude}} 2\pi|u|\,|F(u)|$

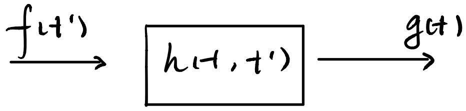


(high-Pass filter)

$\rightarrow$ Integration

$\mathcal{F}\left[\int_{-\infty}^{t} f(t')\,dt'\right] = \frac{1}{i2\pi u} F(u)$

$\rightarrow$ Apply to LSI/LTI



$$g(t) = \int_{-\infty}^{+\infty} f(t')\, h(t-t')\, dt' = f(t') * h(t')$$

$h(t, t') = h(t-t')$

$\Rightarrow$ $\mathcal{F}[g(t)] = G(t)$

$\mathcal{F}[f(t)] = F(t)$ $\quad \Rightarrow G(u) = F(u)\, H(u)$

$\mathcal{F}[h(t')] = H(u)$

v. $f(t) = \alpha\, e^{i 2\pi u_0 t}$ $\qquad G(u) = \alpha\, \delta(u-u_0)\, H(u_0)$

$F(u) = \alpha\, \delta(u-u_0)$ $\qquad g(t) = \alpha\, H(u_0)\, e^{i 2\pi u_0 t}$

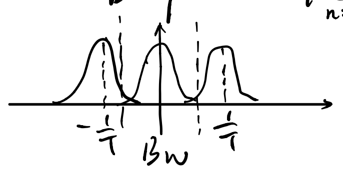
$\rightarrow$ Discrete - Time FT (DTFT)

v. Original continuous functions $f(t)$

v. $f_s(t) = f(t) \cdot \sum_{n=-\infty}^{+\infty} \delta(t-nT)$

v. $\text{Comb}\left(\frac{t}{T}\right) \triangleq T \sum_{n=-\infty}^{+\infty} \delta(t-nT)$

$$\mathcal{F}[f_s(t)] = F(u) * \frac{1}{T} \sum_{n=-\infty}^{+\infty} \delta\left(u - \frac{n}{T}\right) = \frac{1}{T} \sum_{n=-\infty}^{+\infty} F\left(u - \frac{n}{T}\right)$$



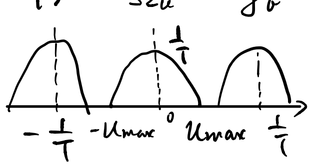
$\longrightarrow$ Nyquist - Shannon Sampling Theorem

If $f(t)$ is band limited i.e. $F(u) = 0$, $|u| > \Omega_u$

then if I sample $f(t)$ using T, s.t. $T < \frac{1}{\Omega_u}$

Then I recover $f(t)$ from discrete samples (uniform)

$T_s = \frac{1}{\Omega_u}$ (Nyquist Sampling rate)



$F'(u) = T\, F_s(u)\, \text{rect}(Tu)$ $\quad \left[-\frac{1}{2T}, \frac{1}{2T}\right]$

$$\downarrow$$

$$f'(t) = T f_s(t) * \frac{1}{T} \text{sinc}\left(\frac{t}{T}\right)$$

$$= \sum f(nT) \left[\delta(t-nT) * \text{sinc}\left(\frac{t}{T}\right)\right]$$

$$= \sum f(nT)\, \text{sinc}\left(\frac{t-nT}{T}\right)$$

$f_s(t) = \sum f(nT)\, \delta(t-nT)$

$\mathcal{F}[f_s(t)] = \sum_{n=-\infty}^{+\infty} f(nT)\,\mathcal{F}[\delta(t-nT)] = \sum f(nT)\,e^{-i2\pi u n T}$

Define $w = 2\pi u \underbrace{T}_{\text{sampling rate}}$ $\quad u \in [-\frac{1}{2T}, \frac{1}{2T}] \Rightarrow w \in [-\pi, \pi]$

$$F(w) = \sum_{n=-\infty}^{+\infty} f[n]\, e^{-inw} \qquad w \in [-\pi, \pi]$$

$$f[n] = \int_{-\pi}^{\pi} F(w)\, e^{iwn}\, dw \qquad \text{Inverse DTFT}$$

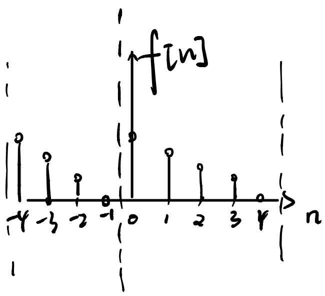
→ DFT (Discrete Fourier Transfrom)

$w \in [-\pi, \pi]$ for main period, if we sample N values in $[-\pi, \pi]$

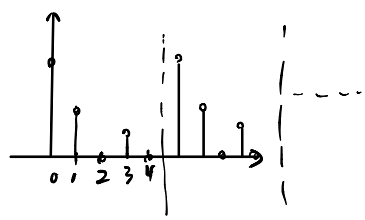
we sample at $w_k = \frac{2\pi}{N} k$

$$F(w_k) = \sum_{n=-\infty}^{+\infty} f[n]\, e^{-i\frac{2\pi}{N}kn} \qquad F(k) = \frac{1}{N}\sum_{n=0}^{N-1} f[n]\, e^{-i\frac{2\pi}{N}kn} \quad \text{DFT}$$

$$f[n] = \sum_{k=0}^{N-1} F(k)\, e^{i\frac{2\pi}{N}kn} \quad \text{Inverse DFT}$$



$\xrightarrow{\text{DFT}}$



Is suffice to only deal with samples within $[0, N-1]$

→ Matrix formulation of DFT

$$F[k] = \frac{1}{N}\sum_{n=0}^{N-1} f[n]\, e^{-i\frac{2\pi}{N}kn}$$

Define:

$$\vec{y} = \begin{bmatrix} F[0] \\ F[1] \\ \vdots \\ F[N-1] \end{bmatrix} \underset{\text{IDFT}}{\overset{\text{DFT}}{\rightleftharpoons}} \vec{f} = \begin{bmatrix} f[0] \\ f[1] \\ \vdots \\ f[N-1] \end{bmatrix} \qquad \vec{y} = W\vec{f}$$

The lth row of $W \triangleq \frac{1}{N}[1, e^{-i\frac{2\pi}{N}l}, e^{-i\frac{2\pi}{N}\cdot 2l}, \cdots, e^{-i\frac{2\pi}{N}l(N-1)}]$

$W_{lp} = e^{-i\frac{2\pi}{N}lp}$

$$W = \frac{1}{N}\begin{array}{c} \left[\begin{array}{ccccc} 1, & 1, & 1, & \cdots\cdots & 1 \\ 1, & e^{-i\frac{2\pi}{N}} & e^{-i\frac{2\pi}{N}\cdot 2}, & \cdots\cdots & e^{-i\frac{2\pi}{N}(N-1)} \\ \vdots & & & & \\ 1, & e^{-\frac{2\pi}{N}k}, & e^{-\frac{2\pi}{N}2k}, & \cdots\cdots, & e^{-i\frac{2\pi}{N}k(N-1)} \\ \vdots & & & & \\ 1, & e^{-\frac{2\pi}{N}k(N-1)}, & & \cdots\cdots, & e^{-i\frac{2\pi}{N}(N-1)^2} \end{array}\right]_{N\times N} \end{array} \begin{array}{l} k=0 \\ k=1 \\ \\ k \\ \\ k=N-1 \end{array}$$

$W = [\vec{q}_0, \vec{q}_1, \cdots, q_{N-1}]_{N\times N}$ $\quad \vec{y} = W\vec{f} = [q_0, q_1, \cdots q_{N-1}]\begin{bmatrix} f[0] \\ f[1] \\ \vdots \\ f[N-1] \end{bmatrix}$

$\vec{y} = \sum_{n=0}^{N-1} f[n]\,\vec{q}_n$ (basis decomposition form)

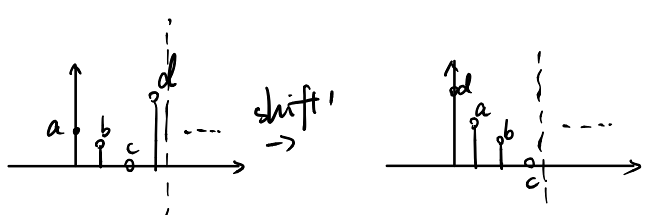
$$\langle \vec{q}_l, \vec{q}_k \rangle = \begin{cases} 0 & l \neq k \\ \frac{1}{N} & l = k \end{cases} \quad \text{(orthogonal)}$$

$$\vec{q}_l = \frac{1}{N}\begin{bmatrix} 1 \\ e^{-i\frac{2\pi}{N}l} \\ e^{-i\frac{2\pi}{N}l2} \\ \vdots \\ e^{-i\frac{2\pi}{N}l(N-1)} \end{bmatrix}$$

Parseval Theorem. concerning energy difference between $f[n]$ & $F[k]$

$$\begin{aligned} \|\vec{y}\|_2^2 &= \|W\vec{f}\|_2^2 \\ &= \langle W\vec{f}, W\vec{f} \rangle = \vec{f}^H W^H W f \qquad W^H W = \begin{bmatrix} \frac{1}{N} & & \\ & \frac{1}{N} & \\ & & \ddots \ \frac{1}{N} \end{bmatrix} \\ &= \frac{1}{N}\vec{f}^H \vec{f} = \frac{1}{N}\|\vec{f}\|_2^2 \end{aligned}$$

Properties of DFT



→ linearity

→ Time shift (circular shift)

Mathematically

$\{x_{\langle n-m\rangle_N}\}_0^{N-1}$

shift by m

$\{x_n\}_{n=0}^{N-1} \xrightarrow{DFT} \{X_k\}_{k=0}^{N-1}$ $\qquad$ $\{x_{\langle n-m \rangle_N}\}_{n=0}^{N-1} \xrightarrow{DFT} \{X_k e^{-i\frac{2\pi}{N}km}\}_{k=0}^{N-1}$

Parseval theorm

$$\underbrace{N \sum_{k=0}^{N-1} |\overset{F_k}{X_k}|^2}_{} = \sum_{n=0}^{N-1} |\overset{f_n}{x_n}|^2$$

l_2-norm square of $\qquad$ l_2-norm of f

$$\vec{y} = \begin{bmatrix} F(0) \\ F(1) \\ \vdots \\ F(N-1) \end{bmatrix} \quad 及 \quad \vec{f} = \begin{bmatrix} \vdots \\ \vdots \end{bmatrix}$$

→ Circular Convolution

Two sequences $\{h_n\}_{n=0}^{N-1}$ $\{f_n\}_{n=0}^{N-1}$

$$g = h \circledast f \triangleq \sum_{m=0}^{N-1} f[m]\, h[\langle n-m \rangle_N]$$

Note: ① Both Sequence should have the same length.

② $len(h_n) = N_1$, $len(f_n) = N_2$

→ zeropad to both sequences to $N = \max\{N_1, N_2\}$

$h \circledast f[n] \xrightarrow{DFT} \{H_k F_k\}_{k=0}^{N-1}$ multipulcation

circular convolution

⇒ linear convolution : $\vec{h}, \vec{f}$ $\qquad$ $\vec{g} = \underbrace{H(\vec{h})}_{} \cdot \vec{f}$

Toeplitz Matrix

$$g[0] = \sum_{m=0}^{N-1} f_1[\langle -m \rangle_N] f_2[m] = f_1[0] f_2[0] + f_1[N-1] f_2[1] + \cdots + f_1[1] f_2[N-1]$$

$$g[1] = \sum_{m=0}^{N-1} f_1[\langle 1-m \rangle_N] f_2[m] = f_1[1] f_2[0] + f_1[0] f_2[1] + \cdots + f_1[2] f_2[N-1]$$

$\vdots$

$$g[N-1] = \sum_{m=0}^{N-1} f_1[\langle N-1-m \rangle_N] f_2[m] = f_1[N-1] f_2[0] + f_1[N-2] f_2[1] + \cdots + f_1[0] f_2[N-1]$$

$$\underset{\vec{y}}{\begin{bmatrix} g[0] \\ g[1] \\ g[2] \\ \vdots \\ g[N-1] \end{bmatrix}} = \underset{\vec{H}}{\begin{bmatrix} f_1[0] & f_1[N-1] & f_1[N-2] & \cdots & f_1[1] \\ f_1[1] & f_1[0] & f_1[N-1] & \cdots & f_1[2] \\ f_1[2] & f_1[1] & f_1[0] & \cdots & f_1[3] \\ \vdots \\ f_1[N-1] & f_1[N-2] & f_1[N-3] & \cdots & f_1[0] \end{bmatrix}_{N \times N}} \underset{\vec{f}}{\begin{bmatrix} f_2[0] \\ f_2[1] \\ \vdots \\ f_2[N-1] \end{bmatrix}}$$

Cricular matrix

$$H = Q^{-1} \Lambda Q$$

Diagonal will be DFT of $\{f_1\}$

$\rightarrow$ connection to linear convolution

$\Rightarrow f_1 * f_2 = g[n]$ length of $\vec{g} = N+M-1$

$\Rightarrow$ pad $\{f_1\}$ to $N+M-1$ with 0's $\Rightarrow \{f_1'\}$

pad $\{f_2\}$ to $N+M-1$ with 0's $\Rightarrow \{f_2'\}$

$f_1' \circledast f_2' = f_1 * f_2$

From 1D to 2D to MD

1D : $F(k) = \frac{1}{N}\sum_{n=0}^{N-1} f[n]\, e^{-i\frac{2\pi}{N}nk}$

2D : $F[k_1,k_2] = \frac{1}{N_1 N_2}\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1} f[n_1,n_2]\, e^{-i\frac{2\pi}{N}n_1k_1}\, e^{-i\frac{2\pi}{N}n_2k_2}$

for each n_2, perform DFT with n_1 (first dimension)

DFT along the second dimension

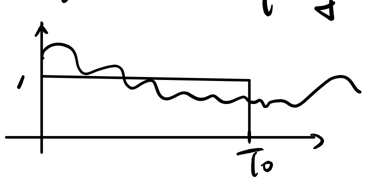
Frequency resolution

$\Rightarrow$ Assumptions : ① $f(t) \in R$ bandlimited

② Generete samples $\{f_s(t)\}$ at Nyquist rate

effects of sampling

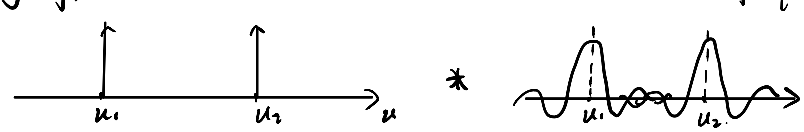
Finite sampling $f_s'(t) = f_s(t)W(t)$



$W(t) = \begin{cases} 1 & t \le T_0 \\ 0 & t > T_0 \end{cases} = rect\left(\frac{t-\frac{T_0}{2}}{T_0}\right)$

$\mathcal{F}[f_s'(t)] = \mathcal{F}[f_s(t)] * \mathcal{F}\left[rect\left(\frac{t-\frac{T_0}{2}}{T_0}\right)\right] \rightarrow sinc(T_0u)e^{-i2\pi u\frac{T_0}{2}}$

$\mathcal{F}[f_s(t)] \overset{\text{example}}{=\!=} \delta(u-u_1) + \delta(u-u_2) \Rightarrow$ 2 frequency components.



ideal Observe

$|u_1 - u_2| > \frac{1}{T_0}$ requirement $\rightarrow$ Frequency ability

$T_0 > \frac{1}{|u_1-u_2|} = \frac{1}{\Delta u}$

The ability to distinguish two close freq. component

T: sampling rate

$NT > \frac{1}{\Delta u}$

$\Updownarrow$

T_0

short - Time Fourier transform (ST - FT)

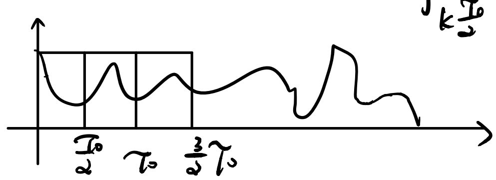
v. Tracking variations of difference frequency components.

v. Recall FT

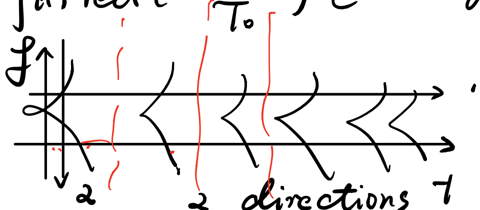
$$F(u) = \int_{-\infty}^{+\infty} f(t) e^{-i2\pi ut} dt$$

truncate

$$F(u, T_0) = \int_{k\frac{T_0}{2} - \frac{T_0}{2}}^{k\frac{T_0}{2} + \frac{T_0}{2}} f(t) \, \mathrm{rect}\left(\frac{t - k\frac{T_0}{2}}{T_0}\right) e^{-i2\pi ut} dt$$



$\rightarrow$



Tradeoffs: ① higher frequency resolution $T_0 \uparrow$ but lose temporal resolution

② higher temporal resolution. $T_0 \downarrow$ lose f resolution

$\longrightarrow$ Symmetric Properties of FT

v For any function $f(t) = f_o(t) + f_e(t)$ (odd, even) (e.g. $f_e(t) = \frac{f(t) + f(-t)}{2}$)

$$F(u) = \int f(t) e^{-i2\pi ut} dt = \int f_e(t) e^{-i2\pi ut} dt + \int f_o(t) e^{-i2\pi ut} dt$$

$\cos(2\pi ut) - i\sin(2\pi ut)$

$$= \int f_e(t) \cos(2\pi ut) dt - i \int f_o(t) \sin(2\pi ut) dt$$

$$= 2\int_0^{\infty} f_e(t) \cos(2\pi ut) dt - 2i \int_0^{\infty} f_o(t) \sin(2\pi ut)$$

① $f(t)$ real even: $F(u) = 2\int_0^{\infty} f_e(t) \cos(2\pi ut) dt$

② $f(t)$ real odd: $F(u) = -2i \int_0^{\infty} f_o(t) \sin(2\pi ut) dt$

If $f(t)$ is real $\mathcal{F}[f(u)] = \mathcal{F}(u)$ $F(-u) = F^*(u)$ (Complex symmetric)

System and Signal representation

$f[n]$ (Discretized object function) $\longrightarrow$ $\boxed{\text{System } h[n]}$ $\longrightarrow$ $g[n]$ (Discretized data being generated)

→ Put DFT in this context

$g[n] = [f \circledast h][n] \longleftrightarrow G(k) = F(k) \cdot H(k)$

$\vec{g} = H\vec{f}$ (view matrix as an operation)

H: a circulant matrix formed from $\{h[n]\}_{n=0}^{N-1}$

$H = Q\Lambda Q^{-1}$ Q^{-1}: DFT Matrix Q: Inverse DFT Matrix

$$Q = \frac{1}{\sqrt{N}}\begin{bmatrix} 1, & 1, & \cdots & , 1 \\ 1, & e^{i\frac{2\pi}{N}}, & \cdots & e^{\frac{i2\pi(N-1)}{N}} \\ \vdots & & & \\ 1, & e^{\frac{i2\pi(N-1)}{N}}, & \cdots & e^{\frac{i2\pi(N-1)^2}{N}} \end{bmatrix}$$

$Q^{-1} = Q^H$ Q is an orthogonal Matrix

$\forall \circledast \iff HQ = Q\Lambda$ $\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{bmatrix}$ diagonal matrix

$$\vec{g} = H\vec{f} \Rightarrow \vec{g} = \underbrace{Q \underbrace{\Lambda \underbrace{Q^{-1}\vec{f}}_{\text{DFT}}}}\ $$

$Q^{-1}\vec{f}$: DFT (Discrete Fourier coefficients)

Λ: scales Fourier coefficients by $\{\lambda_i\}_{i=1}^{N}$

Q: synthesize an output signal using the scaled coefficients with Inverse DFT

$Q = [\vec{q}_1 \cdots \vec{q}_2]_{N\times N}$ $\vec{q}_l = \begin{bmatrix} e^{i\frac{2\pi}{N}l} \\ \vdots \\ e^{i\frac{2\pi}{N}(N-1)} \end{bmatrix}$ lth column of Q

DFT basis for representing a vector

A general linear system

$\vec{g} = H\vec{f}$

→ singular value decomposition (SVD)

Any $H \in R^{M\times N}$

Theorm: $H = USV^T$

∨. $U \in R^{M\times M}$ & $U^TU = I$ (orthogonal)

∨. $V \in R^{N\times N}$ & $V^TV = I$ (orthogonal)

∨. $S \in R^{M\times N}$

$$S = \begin{bmatrix} \sigma_1 & & & \phi \\ & \ddots & & \\ & & \sigma_R & \\ & \phi & & \phi \end{bmatrix}$$

e.g. $M < N$

$R < \min\{M, N\}$

$$S = \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & \phi \\ & & \sigma_M & \end{array}\right]$$

$M\times M$

singular value of H

R is the nonzero signular values of H

$R = \text{rank}(H)$

Definitions:

$\Rightarrow$ Column space of H: span{columns of H}

$\Updownarrow$

Column space of U — combinations of coef (SV^T) and U

Row space of H — span{row of H}

$\Updownarrow$

Row space of V^T

null space of H: $\{\vec{v};\ H\vec{v}=0\}$

rank $\triangleq$ nonzero sigular values of H (S)

$\Downarrow$

Dim (column space) Dim (row space)

$\rightarrow$ Reduced form of SVD

$H = USV^T$ $\qquad R \le \min\{M, N\}$

$\Rightarrow \underbrace{U'}_{M\times R}\underbrace{S'}_{R\times R}\underbrace{V'^T}_{R\times N}$

$H = USV^T \qquad V = [\vec{v}_1, \cdots, \vec{v}_N]_{N\times N}$

$\{v_i\}_{i=1}^N$: a set of basis of R^N

$\vec{g} = H\vec{f} \qquad \vec{f} \in {}^{N\times 1}$

$\vec{f} = \sum_{k=1}^N \alpha_k \vec{v}_k$

$$\vec{g} = USV^T \sum_{i=1}^N \alpha_k \vec{v}_k = USV^T V\vec{\alpha} = US\vec{\alpha}$$

$$S\vec{\alpha} = \begin{bmatrix} \sigma_1\alpha_1 \\ \sigma_2\alpha_2 \\ \vdots \\ \sigma_R\alpha_R \\ \phi \end{bmatrix}$$

$$\vec{g} = US\vec{\alpha} = [\vec{u}_1, \cdots, \vec{u}_M] \begin{bmatrix} \sigma_1\alpha_1 \\ \sigma_2\alpha_2 \\ \vdots \\ \sigma_R\alpha_R \\ \phi \end{bmatrix} = \sum_{i=1}^R \sigma_i \alpha_i \vec{u}_i$$

synthesize the output using $\{\vec{u}_i\}_{i=1}^N$

$\vec{g} = H\vec{f}$

$$\vec{f} = \underbrace{\sum_{k=1}^R \alpha_k \vec{v}_k}_{f_1} + \underbrace{\sum_{k=R+1}^N \alpha_k \vec{v}_k}_{f_2}$$

$R \triangleq \mathrm{rank}(H)$

$$H\vec{f} = \sum_{i=1}^{\min(M,N)} \sigma_i\alpha_i \vec{u}_i = \underbrace{\sum_{i=1}^R \sigma_i\alpha_i \vec{u}_i}_{H\vec{f}_1} + \underbrace{\sum_{i=R+1}^{\min(M,N)} \sigma_i\alpha_i \vec{u}_i}_{H\vec{f}_2}$$

$H\vec{f} = H\vec{f}_1 + H\vec{f}_2$ $\quad$ $\vec{g}_2 = Hf_2 = USV^T f_2 = 0$

All $\vec{f}_2 = \sum_{k=R+1}^{\min\{M,N\}}$ is in the null space

$\Updownarrow$

All $\vec{f}_2 \neq 0$ can't be measured by H

or $\vec{f} = f_1 + f_2$. Any f with the same $\vec{f}_1$ will operate the same $\vec{g}$

Case ① $\quad \vec{g} = H\vec{f} \quad H \in R^{M\times N} \quad \mathrm{rank}(H) = N \quad (M \geq N)$

only element in the null space of H is 0

Case ② $\quad M < N \quad \mathrm{rank}(H) = R \leq M$

$\vec{g} = [\quad]_{M\times N} \vec{f}$ $\quad$ H has a null space $(N-R)$ dimension

→ SVD as dimensionality reduction tool

Given $X \in R^{M\times N}$ $\quad \begin{bmatrix} \vec{x}_1^T \\ \vec{x}_2^T \\ \vdots \\ \vec{x}_m^T \end{bmatrix}$ → different measurements of a variable.

low rank approximation of X

Goal: Find X_L where

$\mathrm{rank}(X_L) \leq R$

Approach: $\quad \min \|X - X_L\|_F^2$

original $\quad$ residual of approximation

$\mathrm{rank}(X_L) \leq R$

$X_L = \underset{M\times R}{U} \underset{R\times R}{S} \underset{R\times N}{V^T}$

above

U: orthogonal (different columns)

V: orthogonal (different rows)

S: diagonal

→ System inversion (Recovery of Object)

$\vec{g} = H\vec{f}$ $\quad \vec{f} \to \boxed{H} \to g \to \boxed{R} \to \hat{f} \approx f$

known. known. $\quad$ signal recovery (linear inverse problem)

recovery

Midterm 3 & 5

1. what is linear system propoties.

what is LTI (LSI) system

→ Conv Corr

$f(t)$ $g(t)$

$f[n]$, $g[n]$

Conv connect LTI

→ 4 Fourier Transform.

FS FT DTFT DFT propoties

linearity shifting conv $\delta(t)$

→ Matrix represntation

① linear System (LSI , general LS

② FT (DFT)

③ Conv circulen Conv

→ Signal decomposition

$f(t) = \sum \alpha_k b_k(t)$ $f[n] = \sum_{k=1}^{N} \alpha_k b_k[n]$

orthonomarl basis. linear independence.

→ Sampling Theorm.

what why how SVD

case ① $H \in R^{N \times N}$

case ② $H \in R^{M \times N}$ $(M < N)$

$g_{M \times 1} = [h_1, h_2, \cdots h_N] f_{N \times 1}$ (underdetermined)

$h \in R^{M \times 1}$ Maximum possible $\text{rank}(H) = M < N$

$H\vec{v} = \vec{0}$: There are many possible v's $v \in R^N$

$\Updownarrow$

H has nontrivial null space

$\Rightarrow$ Assume $f^{(1)}$ is a solution of $\vec{g} = Hf^{(1)}$

Any $f' = f^{(1)} + \alpha\vec{v}$ ($\vec{v}$ from the null space of H)

is also a solution of $\vec{g} = H\vec{f}$

Minimum norm solution ($\vec{q} = H\vec{f}$ $\text{rank}(H) = M$ $M < N$

$$f_{MN} = \underset{\{\vec{f},\, H\vec{f} = \vec{g}\}}{\arg\min} \|\vec{f}\| \qquad f_{MN} = H^T(HH^T)^{-1}\vec{g}$$

$$\vec{f} \rightarrow \boxed{H} \rightarrow \vec{g} \rightarrow \boxed{G} \ \hat{f}$$

case ③ $H \in R^{M \times N}$ $M > N$

$$g_{M \times 1} = \begin{bmatrix} h_1, h_2 \cdots h_N \\ h_i \in R^{M \times 1} \end{bmatrix} f_{N \times 1}$$

$\text{rank}(H) = N$

over-determined

$r = Hf - g$ ideally $r = 0$

$$\hat{f} = \underset{f \in R^N}{\arg\min} \underbrace{\|Hf - g\|_2^2}_{\text{least squares}} \qquad \hat{f} \triangleq f_{LS}$$

$f_{LS} = (H^T H)^{-1} H^T g$ $(M > N)$

$f_{MN} = H^T (HH^T)^{-1} g$ $(M < N)$

$$H^T H = VS^T S V^T$$
$$= VS^2V^T$$
$$= (VS^2V^T)^{-1} = VS^{-2}V^T$$

Stochastic Modeling of data

Example:

$f(t) \rightarrow \underset{\text{(System)}}{H} \rightarrow g(t) \rightarrow \oplus \rightarrow \underline{g'(t)}$

object: $f(t)$ $h(t)$ data: $g(t)$ $n(t)$ noise   observed: $g'(t)$

For a given t, evaluate $f(t)$ or $g(t)$

LSI / LTI

$$g'(t) = \underbrace{\int_{-\infty}^{+\infty} f(t')\, h(t-t')\, dt}_{\text{ideal measurements}} + \underbrace{n(t)}_{\text{random effects}}$$

How can we model the random effects in the data?

→ Random Process

Example. $X(t) = A\cos(wt + \theta)$ $\begin{cases} w: \text{frequency.} \\ \theta: \text{phase} \end{cases}$

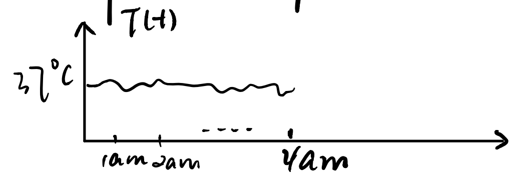
A. random variable $\sim N(0,1)$

For a specific $a \sim N(0,1)$, you can plot a specific waveform each one is called "sampled curve".

$\boxed{X(t_0) = \underset{\text{R.V.}}{A}\cos(wt_0 + \theta)} \in$ R.V.

$\bar{X}(t) \triangleq \frac{1}{N}\sum_{n=1}^{N} x_n(t)$ ensemble mean of random process

Example: Temperature accross the day.



Day 1

Day 2

$T(2am) \sim N(30.5°C,\ 0.8°C$ $T(3am)$ $T(4am)$ --- $T(3pm)$ ---

$X(t)$: is a random process, for each t_0. $X(t_0)$ is a random variable (continuous-time R.P.)

$x(n) \triangleq X(n\,T)$ Discrete-time R.P.

Modeling the random process

✓ Joint PDF

Given $X(t)$, a random process

Sample $X(t)$ at a set of indices. $\{t_i\}_{i=1}^{N}$

$[X(t_1), X(t_2), \cdots X(t_N)]^T = \vec{X}$

$P_{\vec{X}}(x(t_1), x(t_2), \cdots, x(t_N))$: Joint PDF

$$\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_N} P_{\vec{X}}(x(t_1), x(t_2), \cdots x(t_N))\,dx_1\,dx_2\cdots dx_N$$

$$= P(x(t_1)\le x_1,\ x(t_2)\le x_2, \cdots\ x(t_N)\le x_N)$$

Marginal distribution

$$\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} P_{\vec{X}}(x(t_1), x(t_2), \cdots x(t_N))\,dx_1\,dx_2\cdots dx_{k-1}\,dx_{k+1}\cdots dx_N$$

$$= P(x(t_k))$$

Example: multivariant Gaussian distribution

$[X_1, \cdots, X_N]^T$ $\quad P_{\vec{X}}(x_1, x_2, \cdots x_N) = (2\pi)^N (\det K)]^{-\frac{1}{2}}$

$$\exp\left(-\tfrac{1}{2}(\vec{x}-\vec{\mu})^T K^{-1} (\vec{x}-\vec{\mu})\right)$$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \quad \vec{\mu} = \begin{bmatrix} E[X_1] \\ \vdots \\ E[X_N] \end{bmatrix}$$

K: covariance matrix

$[X(t_1), X(t_2), X(t_3), \cdots, X(t_N)]^T \in R^{N\times 1}$

$\mathrm{Var}[X(t_i)] = E[(X(t_i) - \mu_i)^2]$

$$\begin{cases} \mathrm{Cov}[X(t_i)\,X(t_j)] = E[(X(t_i)-\mu_i)(X(t_j)-\mu_j)] \\ \mathrm{Corr}[X(t_i)\,X(t_j)] = E[X(t_i)\,X(t_j)] \end{cases}$$

Stationary R.P. (strict-sense R.P.)

$p_{\vec{X}_1}(x(t_1), x(t_2), \cdots x(t_N)) = p_{\vec{X}_2}(x(t_1+\tau), x(t_2+\tau), \cdots, x(t_N+\tau))$

$p_x(x(t_1)) = p_x(x(t_1+\tau))$ PDF is shift invariant

- Wide-sense stationary.

$\Rightarrow E[X(t_1)] = E[X(t_1+\tau)] = \mu \in$ constant.

Generally $X(t)$ $\quad E[X(t)] = \mu(t)$

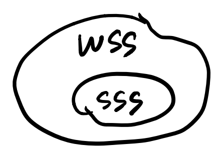
$\Rightarrow Cov[X(t_1), X(t_2)] = Cov[X(t_1+\tau), X(t_2+\tau)]$

only depends ~~of~~ on the differences of t_1 & t_2.

$\Downarrow$

$\Rightarrow Var[X(t_1)] = Var[X(t_1+\tau)] = C$ (constant)

Then $X(t)$ is a WSS R.P.



Example.

$$X(t) = \begin{cases} \sin(t) & \text{prob } \frac{1}{4} \\ -\sin(t) & \text{prob } \frac{1}{4} \\ \cos(t) & \text{prob } \frac{1}{4} \\ -\cos(t) & \text{prob } \frac{1}{4} \end{cases}$$

$E[X(t)] = 0$ ✓

$Cov[X(t_1), X(t_2)] = E[X(t_1)X(t_2)]$

$= \sin(t_1)\sin(t_2)\cdot\frac{1}{16} - \sin(t_1)\sin(t_2)\cdot\frac{1}{16}$ ......

$= \frac{1}{2}\cos(t_2 - t_1)$ ✓ $\quad$ WSS ✓

$$X(0) = \begin{cases} 0 & \frac{1}{4} \\ 0 & \frac{1}{4} \\ 1 & \frac{1}{4} \\ -1 & \frac{1}{4} \end{cases} \quad \left.\right\} \frac{1}{2} \qquad X(\tfrac{1}{4}) = \begin{cases} \frac{\sqrt{2}}{2} & \frac{1}{4} \\ -\frac{\sqrt{2}}{2} & \frac{1}{4} \\ \frac{\sqrt{2}}{2} & \frac{1}{4} \\ -\frac{\sqrt{2}}{2} & \frac{1}{4} \end{cases}$$

SSS ✗

- Random process $X(t)$: A sequence of Random variable

t index variable $t \in R$

$x(n)$: n is a integar index

Given a set of index $\{t_i\}$

$\{X(t_1), X(t_2), \cdots, X(t_N)\}$ Modeled by the joint PDF

⇒ Auto-correlation function (ACF) (for WSS)

$E[X(t)X(t-\tau)] = \phi_X(\tau) = \phi_X(-\tau)$

⇒ Ergodic Random Process (WSS)

WSS $X(t)$

$$\underbrace{E[X]}_{\text{Ensemble average}} = \lim_{T_t \to \infty} \frac{1}{T_t} \underbrace{\int_0^{T_t} x(t)\,dt}_{\text{Time average}}$$

Then $X(t)$ is an Ergodic Process

$E[X(n)] = \lim_{T\to\infty} \frac{1}{T} \sum_{k=1}^{T} x_i(kT)$

→ Example $X(t)$ is a Random Process and it's also i.i.d.

For any $\{t_i\}$ $[X(t_1), X(t_2), \cdots X(t_N)]^T$

For a sampled curve $x(t_1)\ x(t_2) \cdots x(t_N)$

→ Gaussian Process

Definition $\{t_i\}_{i=1}^{N}$ $[X(t_1), X(t_2) \cdots X(t_N)]^T$

$$\begin{bmatrix} X(t_1) \\ X(t_2) \\ \vdots \\ X(t_N) \end{bmatrix} \sim N(\vec{\mu}, K) \begin{bmatrix} \text{Joint multivariant} \\ \text{Gussian distribution} \end{bmatrix}$$

$$\vec{\mu} = \begin{matrix} E[X(t_1)] & \mu_1 \\ E[X(t_2)] & \mu_2 \\ \vdots & \\ E[X(t_N)] & \mu_N \end{matrix}$$

$(\vec{X} - \vec{\mu}) \in R^{N\times 1}$

$E[(\vec{X}-\vec{\mu})(\vec{X}-\vec{\mu})^T]$

$$E\begin{bmatrix} \begin{bmatrix} X(t_1)-\mu_1 \\ X(t_2)-\mu_2 \\ \vdots \\ X(t_N)-\mu_N \end{bmatrix} [X(t_1)-\mu_1, X(t_2)-\mu_2 \cdots \end{bmatrix}$$

$$= \begin{bmatrix} E[(X(t_1)-\mu_1)(X(t_2)-\mu_2)], E[(X(t_2)-\mu_2)(X(t_2)-\mu_2)] \cdots \\ \underbrace{E[(X(t_2)-\mu_2)(X(t_1)-\mu_1)]}_{K_X(t_1,t_2)}, E[(X(t_2)-\mu_2)(X(t_2)-\mu_2)] \cdots \\ \vdots \qquad \ddots \\ \qquad\qquad \text{rank}=1 \end{bmatrix}$$

Var = constant

$$K = \begin{bmatrix} K_X(t_1,t_1) & K_X(t_1,t_2) & K_X(t_1,t_3) & \cdots & K_X(t_1,t_N) \\ K_X(t_2,t_1) & K_X(t_2,t_2) & \cdots & & \cdots \\ K_X(t_N,t_1) & \cdots & & & K_X(t_N,t_N) \end{bmatrix}$$ Symmetric

$$P_{\vec{X}}(x_1,\cdots,x_N) = [(2\pi)^N \det(K)]^{-\frac{1}{2}} \exp[-\tfrac{1}{2}(x-\vec{\mu})^T K^{-1}(x-\vec{\mu})]$$

mean vector, cov Matrix

$\vec{X} = X_1$ $K = var(X_1)$

→ White Gaussian noise (process)

$[X(t_1), X(t_2), \cdots, X(t_N)]^T \sim N(0, \sigma^2 I_N)$

① $\mu = 0$ ② $var(X_{t_i}) = \sigma^2$ ③ any two $X(t_i), X(t_j)$ are independent

④ $\phi_X(\tau) = \sigma^2 \delta(\tau)$ $\delta(\tau) = \begin{cases} 1 & \tau = 0 \\ 0 & \text{otherwise} \end{cases}$

Power spectral density (PSD)

$X(t)$ $X(u) = \int_{-\infty}^{+\infty} X(t) e^{-i2\pi ut} dt$ (Δ)

Problem: ① $|X(t)|$ can be arbitary large ② $X(u)$ is also a R.P.

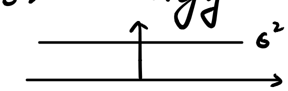
Instead of defining FT of $X(t)$, we define FT of $\phi_X(\tau)$ WSS

PSD: $S_X(u) = \int_{-\infty}^{+\infty} \phi_X(\tau) e^{-i2\pi u\tau} d\tau$ $\phi_X(\tau) = \int_{-\infty}^{+\infty} S_X(u) e^{i2\pi u\tau} du$

Wiener - khinchine Theorm (WKT)

$\phi_X(0) = \int_{-\infty}^{+\infty} S_X(u)\, du \Leftrightarrow var$ (WSS) (energy)

WGN: $\phi_X(\tau) = \sigma^2 \delta(\tau) \xrightarrow{\mathcal{F}} S_X(u) = \sigma^2$



→ Passing Random Process through LTI

$f(t) \rightarrow \boxed{H}_{LTI} \rightarrow g(t)$

R.P. WSS R.P. WSS

$g(t) = \int_{-\infty}^{+\infty} h(t-t') f(t')\, dt'$

$G(u) = F(u) H(u)$

$$\begin{cases} \phi_X(\tau) \xleftrightarrow{\mathcal{F}} S_X(\omega) & \text{input} \\ \phi_Y(\tau) \longleftrightarrow S_Y(\omega) & \text{output} \end{cases} \qquad X(t) \longrightarrow \boxed{\begin{matrix}\text{System} \\ (t)\end{matrix}} \longrightarrow Y(t)$$

$$E[Y(t')\,Y(t'-t)] = \phi_Y(t)$$

$$E\left[\int_{-\infty}^{+\infty} h(a)\,X(t'-a)\,da \int_{-\infty}^{+\infty} h(b)\,X(t'-t-b)\,db\right]$$

$$= E\int db\,h(b)\int h(a)\,X(t'-a)\,X(t'-t-b)\,da = \int db\,h(b)\int_{-\infty}^{+\infty} h(a)\,\underbrace{E[X(t'-a)\,X(t'-t-b)]}_{\phi_X(t+b-a)}\,da$$

$$= \int db\,h(b)\int_{-\infty}^{+\infty} h(a)\,\phi_X(t+b-a)\,da \implies \phi_Y(t) = [h(t) * \phi_X(t)] * h(-t)$$

$$\begin{cases} \mathcal{F}[\phi_Y(t)] = S_Y(\omega) = S_X(\omega)\cdot|H(\omega)|^2 & (\text{using } \mathcal{F}[h(-t)] = H^*(\omega) \\ \phi_Y(\tau) = \phi_X(\tau) * h(\tau) * h(-\tau) \end{cases}$$

→ noise filtering

$$x(t) \longrightarrow \begin{matrix} h(t) \\ LTI \end{matrix} \longrightarrow g(t) \qquad g(t) = \int_{-\infty}^{\infty} h(t')\,x(t-t')\,dt$$

$$g(t) \uparrow \oplus \; n'(t)$$

$$x(t) \longrightarrow h(t) \longrightarrow \oplus \longrightarrow g'(t) \longrightarrow \boxed{F}$$

$$\uparrow n(t) \quad \text{noise R.P.}$$

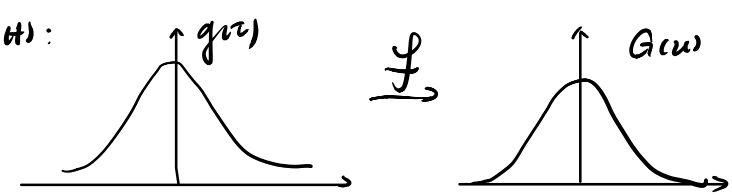
$$g'(t) = g(t) + n(t) \longrightarrow \begin{matrix} F \\ f(t) \,/\, F(\omega) \end{matrix} \longrightarrow \hat{g}(t)$$

Example: $n(t)$: i.i.d. noise (independent & identically distributed

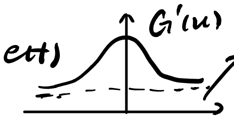
$$E[n(t)] = 0 \qquad E[n(t)\,n(t-\tau)] = \begin{cases} 0 & \tau \neq 0 \\ Var[n(t)] & \tau = 0 \end{cases}$$

$n(t)$: $\phi_N(\tau) = \sigma^2\,\delta(\tau) \xrightarrow{\mathcal{F}} \sigma^2$ (in ω)

signal $g(t)$:



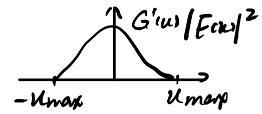
$g'(t) = g(t) + e(t)$

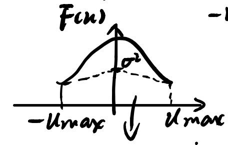


"noise floor"

$\to \boxed{F} \to$

$F(u)$

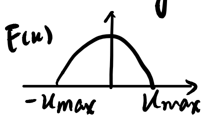


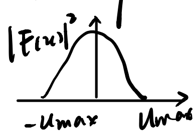


scaled

new noise floor
less than before

You design a low-pass filter





$g(t) \Rightarrow \boxed{F} \to (\Delta)$

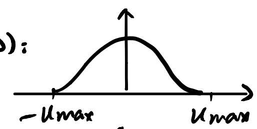
$+$

$e(t) \Rightarrow \boxed{F} \to (*)$

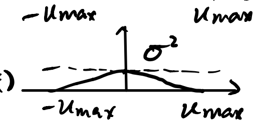
$\shortparallel$

$g'(t) \Rightarrow \boxed{F} \to (\text{Goal})$

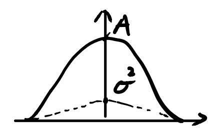
(Δ):



$(*)$



$\oplus$



$\to$ Linear transform of Random vectors

Given a R.P. $X(t)$, $\{t_i\}_{i=1}^{N}$, sample $X(t)$ & get

$\vec{X}(t) = [X(t_1), X(t_2), \cdots, X(t_N)]^T$ Random vector

$\vec{X} \to \boxed{H} \to \vec{Y}$ $\qquad \underset{M\times1}{\vec{Y}} = H \vec{X}_{N\times1}$

System matrix

$$E[\vec{X}] = \begin{bmatrix} E[X(t_1)] \\ E[X(t_2)] \\ \vdots \\ E[X(t_N)] \end{bmatrix} = \mu_x \text{ (mean vector)} \qquad Cov(\vec{X},\vec{X}) \triangleq K_x = E[\underbrace{(\vec{X}-\vec{\mu}_x)}_{N\times1}\underbrace{(\vec{X}-\vec{\mu}_x)^T}_{1\times N}]$$

$$\vec{Y} = H\vec{X} \qquad \begin{cases} E[\vec{Y}] = H E[\vec{X}], \quad E[\vec{X}H^T] = E[\vec{X}]H^T \\ \underset{K_Y}{Cov[\vec{Y},\vec{Y}]} = H \underset{K_X}{Cov(\vec{X},\vec{X})} H^T \end{cases}$$

proof: $K_Y = E[(\vec{Y}-\vec{\mu}_Y)(\vec{Y}-\vec{\mu}_Y)^T] = E[(H\vec{X} - H\vec{\mu}_x)(\vec{X}^T H^T - \vec{\mu}_x^T H^T)]$

$= E[H(\vec{X}-\vec{\mu}_x)(\vec{X}-\vec{\mu}_x)^T H^T] = H\, Cov(\vec{X},\vec{X}) H^T$

$\to$ Moment Generating Function (MGF) and Characteristic function (CCF)

Given a R.V. X $\qquad M_X(s) = E[e^{sX}] = \int_{-\infty}^{+\infty} e^{sx} f(x)\, dx$

$$\frac{dM_X(s)}{ds} = E[e^{sX}X]\Big|_{s=0} = E[X] \qquad \frac{dM_X(s)}{d^2s} = E\left[\frac{d(e^{sX}X)}{ds}\right] = E[e^{sX}X^2]\Big|_{s=0}$$

$$= E[X^2]$$

$$\frac{dM_X(s)}{d^m s}\Big|_{s=0} = E[X^m]$$

Example: $X \sim N(0,1)$ $\quad E[e^{sX}] = \int e^{sx}\frac{1}{\sqrt{2\pi}}e^{-x^2}dx = e^{s^2/2}$

$$\frac{de^{\frac{s^2}{2}}}{ds}\Big|_{s=0} = 0$$

$M_X(s)$ for R.V. X

$$\boxed{\begin{aligned} M_X(u) &= E[e^{-i2\pi uX}] \\ &= \int_{-\infty}^{+\infty} e^{-i2\pi ux}\underbrace{f(x)}_{\text{PDF}}\,dx \end{aligned}} \leftarrow \text{CF} \quad C_X(u)$$

MGF: $M_X(s) = E[e^{sX}]$ $\quad \frac{dM_X(s)}{d^m s}\Big|_{s=0} = E[X^m]$

CF: $s = -i2\pi u$ (or $s = i2\pi u$ or $s = i\Omega$

CF & PDF are FT pairs

Given any two R.V. X & Y

$C_X(u)$: CF of X $\qquad C_Y(u)$: CF of Y $\qquad Z = X+Y$

$$\Pr(Z \le z) = \Pr(X+Y \le Z) = \iint_{X+Y\le Z} f(x,y)\,dx\,dy$$

$$C_Z(u) = E[e^{-i2\pi z}] = E[e^{-i2\pi uX}e^{-i2\pi uY}] = C_X(u)\,C_Y(u)$$

$$\text{PDF}(Z) \triangleq f_Z(z) = f_X(x) * f_Y(y)$$

if $Z = X_1 + X_2 + \cdots + X_N$ $\qquad C_Z(u) = \prod_{i=1}^{N} C_{X_i}(u)$

$$f_Z(z) = f_{X_1}(x_1) * f_{X_2}(x_2) * f_{X_3}(x_3) * \cdots * f_{X_N}(x_N)$$

$\to$ Example $X \sim N(\mu, \sigma^2)$ $\quad Y \sim N(\mu, \sigma^2)$ $\quad$ independent

$Z \triangleq X+Y$ $\qquad E[Z] = 2\mu$ $\qquad P_X(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $\qquad P_Y(y) = \cdots$

$\qquad\qquad\qquad \mathrm{Var}[Z] = 2\sigma^2$

$C_X(u) = \mathcal{F}[P_X(x)]$ $\qquad$ Recall: $\mathcal{F}[e^{-a^2x^2}] = \frac{\sqrt{\pi}}{a}e^{-\pi^2u^2/a^2}$

$$\mathcal{F}[e^{-\frac{x^2}{2\sigma^2}}] = \widehat{\sqrt{2\pi}\sigma}\, e^{-2\pi^2u^2\sigma^2}$$

$$C_Z(u) = e^{-2\pi^2\sigma^2u^2}\cdot e^{-i2\pi\mu u}\cdot e^{-2\pi^2\sigma^2u^2}\cdot e^{-i2\pi\mu u}$$

$$= e^{-4\pi^2\sigma^2u^2}e^{-i4\pi\mu u} \leftarrow \text{Gaussian Distribution}$$

→ Maximum likelihood estimation (Linear transform of random vectors

$f \to \boxed{H} \to \underset{\uparrow \vec{n}}{\oplus} \to \vec{g}$ $\qquad \vec{g} = H\vec{f} + \vec{n}$

A very common model is $\vec{n} \sim N(\vec{\mu}, \vec{k})$ $\quad pdf_N(n) = (2\pi)^{-\frac{N}{2}} (\det k)^{-\frac{1}{2}} \exp(-\frac{1}{2}(\vec{n}-\mu)^T k^{-1} (\vec{n}-\vec{\mu}))$

Q: What's the distribution of data?

$$\begin{cases} E[\vec{g}] = H\vec{f} + \mu \\ Cov[\vec{g}, \vec{g}] = Cov(H\vec{f}+\vec{n}, H\vec{f}+\vec{n}) = Cov(\vec{n}, \vec{n}) = K \end{cases}$$

↙ f is not random

⇓

$$PDF_{\vec{g}}(\vec{x}) = (2\pi)^{-\frac{N}{2}} (\det k)^{-\frac{1}{2}} \exp[-\frac{1}{2}(\vec{x} - H\vec{f})^T k^{-1} (\vec{x} - H\vec{f})]$$

↑ likelihood function

true data has highest possibility $\qquad$ MLE

$$\max_{\hat{f}} (2\pi)^{-\frac{N}{2}} (\det k)^{-\frac{1}{2}} \exp[-\frac{1}{2}(\vec{x} - H\vec{f})^T k^{-1} (\vec{x} - H\vec{f})]$$

↓ log

$$\max_{\hat{f}} A - \frac{1}{2}(\vec{x} - H\vec{f})^T k^{-1} (\vec{x} - H\vec{f}) \underset{\text{i.i.d.}}{\Rightarrow} \min_{\hat{f}} \frac{1}{2\sigma^2} \|\vec{x} - H\vec{f}\|_2^2 \quad \text{LS Solution}$$

WGN ⇕ Gaussian.

Maximum likelihood estimation $\vec{g} = H\vec{f} + \vec{n}$

Preliminary of Statistical Decision (Hypothesis testing)

Assume X is a random variable that represent you a biological parameter you measure.

measurements interpreted as samples drawn from the distribution of X

i.i.d samples of $\{x_i\}$

Estimate some statistics of X from $\{x_i\}_{i=1}$

sample mean: $\bar{x} = \frac{1}{N} \sum x_i$ $\qquad$ sample variance: $S^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2$

Sample mean and sample variance are estimates of true mean (μ) and true variance

→ Sample mean

$E[\bar{x}] = \frac{1}{N} \sum_{i=1}^{N} E[x_i] = \mu = E[X]$ $\quad$ (unbaisedness: $\bar{X}$ is an unbaised estimator of μ

Generally: if $\hat{\theta}$ is an estimator of θ $E[\hat{\theta}] = \theta$, $\hat{\theta}$ is an unbaised estimator of θ

$$Var[\bar{X}] = Var\left[\frac{1}{N}\sum_{n=1}^{N} x_i\right] = \frac{1}{N^2} E\left[\sum_{i=1}^{N} x_i\right] = \frac{1}{N^2} N \sigma^2 = \frac{6^2}{N} \xrightarrow{N\to\infty} 0$$

→ Sample variance: $S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$

$$E[S^2] = \frac{1}{N-1}\sum_{i=1}^{N} E[x_i - \bar{x}]^2 = \frac{1}{N-1}\sum_{i=1}^{N}\{E[x_i^2] - 2E[x_i]\bar{X} + \bar{X}^2\}$$

$$= \frac{1}{N-1}\left\{\sum_{i=1}^{N} E[x_i^2] - 2N\,E[\bar{X}] + E[\bar{X}^2]\right\} = \frac{1}{N-1} E\left\{\sum_{i=1}^{N} x_i - N\bar{X}^2\right\}$$

$$= \frac{1}{N-1}\left\{\sum_{i=1}^{N} E(X_i)^2 - \mu E[\bar{X}^2]\right\}$$

$$E[x_i^2] = Var[x_i] + E^2[x_i] = \sigma^2 + \mu^2 \qquad E[\bar{x}^2] = \frac{1}{N}6^2 + \mu^2$$

$$E[S^2] = \left\{\sum_{i=1}^{N}(6^2 + \mu^2) - N\left[\frac{6^2}{N} + \mu^2\right]\right\} = \frac{1}{N-1}\left\{\sum_{i=1}^{N} 6^2 - 6^2 = \frac{1}{N-1}(N-1)6^2 = 6^2\right.$$

S^2 is unbaised estimator of 6^2

→ Hyposis testing

A single distribution case

- physiological parameter $x \in N(\mu, 6^2)$ healthy population
- New sample / New subject $\hat{x}$ (measurement of sample)

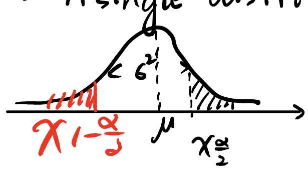
whether $\hat{x} \sim N(\mu, 6^2)$

Hyposis testing { H_0. Null Hyposis $\hat{x}$ belongs to healthy distribution

rejecting the null hyposis | H_1: Competing Hyposis $\hat{x}$ does not belongs to --

Facts: { error will occur
need to characterize the error

→ A single distribution



$$Pr(X \geq x_{\frac{\alpha}{2}}) = \frac{\alpha}{2} \triangleq \frac{\alpha}{2}$$

$$Pr(X \leq x_{1-\frac{\alpha}{2}}) = 1 - Pr(X \geq x_{1-\frac{\alpha}{2}}) = 1 - (1 - \frac{\alpha}{2}) = \frac{\alpha}{2}$$

Decision rule: If $\hat{x} \geq x_{\frac{\alpha}{2}}$ or $\hat{x} \leq x_{1-\frac{\alpha}{2}}$, then $\hat{x}$ does not belong to heathy population (rejecting the null hyposis)

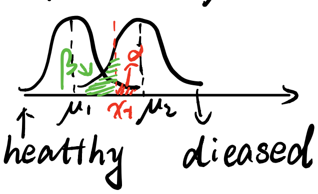
Probobility of error:

$$Pr(X \geq x_{\frac{\alpha}{2}}) + Pr(X \leq x_{1-\frac{\alpha}{2}}) = \alpha \quad \text{(confidence interval)}$$

→ Two distributions (Two populations)

$X \sim N(\mu_1, \sigma^2)$ heathy population

$Y \sim N(\mu_2, \sigma^2)$ dieased population (same distribution of two parameters)



Problem: Given a measurement from a test subject, whether from heathy / dieased

Decision rule: construct a threhold x_t

$$\begin{cases} \text{if } \hat{x} \geq x_t & \text{diseased} \\ \text{if } \hat{x} < x_t & \text{heathy} \end{cases}$$

Errors: $Pr(\hat{x} \geq x_t \mid \text{healthy}) = \alpha$ conditional prob

(Type I error. false positive.

$Pr(\hat{x} \leq x_t \mid \text{diseased}) = \beta$

(Type II error false negative

$X \sim N(\mu, \sigma^2)$ $\quad Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

Threhold $\quad x_\alpha \mid x_{\frac{\alpha}{2}} \longrightarrow \frac{x_\alpha - \mu}{\sigma} = z_\alpha$

$Pr(X \geq x_\alpha) = Pr\left(\frac{X-\mu}{\sigma} \geq \frac{x_\alpha - \mu}{\sigma}\right) = Pr(Z \geq z_\alpha) \leftarrow$ z-score

→ Distinguish means from two populations

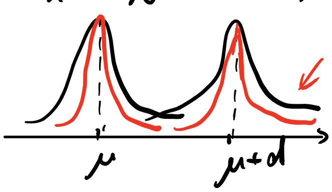
Q: Are the means from two populations you are trying to compare are statistically, significantly different?

Hyposis testing

H_0: the means of X & Y are from same distribution

H_1: - — - - - - - not - - - - - - — —

$\bar{X} = \frac{1}{N}\sum x_i$, $\bar{Y} = \frac{1}{N}\sum y_i$ $\bar{X} \sim N(\mu, \frac{\sigma^2}{N})$ $\bar{Y} \sim N(\mu+d, \frac{\sigma^2}{N})$



Q: How do I choose N

Can I choose a threhold x_t, s.t.

$(z \in N(0,1))$

error prob of $\bar{X}$ to be α $(x_t = x_\alpha)$ $\bar{X} = \frac{\sigma}{\sqrt{N}} z + \mu$

error prob for $\bar{Y}$ to be β $(x_t = x_\beta)$ $\bar{Y} = \frac{\sigma}{\sqrt{N}} z + \mu + d$, x_t

$\frac{\sigma}{\sqrt{N}} z_\alpha + \cancel{\mu} = \frac{\sigma}{\sqrt{N}} z_{1-\beta} + \cancel{\mu} + \beta \Rightarrow \frac{\sigma}{\sqrt{N}}(z_\alpha - z_{1-\beta}) = d \Rightarrow N = \frac{\sigma^2}{d^2}(z_\alpha + z_\beta)^2$

x_α: $Pr(\bar{X} \geq x_\alpha$ | in the eye of the distribution of $\bar{X}) = \alpha$

$Pr(\bar{Y} \leq x_{1-\beta}$ | in the eye of the distribution of $\bar{Y}) = \beta$

α, β is decided by us

e.g. $\alpha = 0.05$, $\beta = 0.05$ $x_\alpha = \frac{\sigma}{\sqrt{N}} z_\alpha + \mu$ z_α: $Pr(z \geq z_\alpha) = \alpha$

$y_{1-\beta} = \frac{\sigma}{\sqrt{N}} z_{1-\beta} + \mu + d$ $z_{1-\beta} = -z_\beta$ $(z \sim N(0,1))$

$x_\alpha = x_{1-\beta} \Rightarrow N = \frac{\sigma^2}{d^2}(z_\alpha + z_\beta)^2$

① $\alpha = \beta$, We want α & β as small as possible $\Rightarrow z_\alpha, z_\beta \uparrow \Rightarrow N\uparrow$

② fixed α, β, $\sigma^2\uparrow \Rightarrow N\uparrow$, given a fixed d.

③ $N \propto \frac{\sigma^2}{d^2}$ e.g. I want to differentiate $\frac{\sigma}{d} = 100$ $N = 10000(z_\alpha + z_\beta)^2$

(precision)

→ Measuring the performance of statistical decision (Binary decision)

⇒ Two population

$X_1 \sim N(\mu_1, \sigma_1^2)$
$X_2 \sim N(\mu_2, \sigma_2^2)$ (Example) X_1: control, healthy
X_2: diseased

Problem: Take a measurement $\hat{x}$ from a new subject.

Decide: whether $\hat{x}$ is heatlhy or not

Strategy: Set a threhold x_t

Decision rule: $\begin{cases} \text{if } \hat{x} \geq x_t, \text{ the person has the disease} \\ \text{if } \hat{x} < x_t, \text{ the person is heatlhy} \end{cases}$

Notations: Test Positive for a disease P

Test Negative for a disease P^c

$\theta_1 \Leftrightarrow$ The subject has the disease W

$\theta_0 \Leftrightarrow$ The subject is heatlhy W^c

Make a decision that is diseased D_1

Make a decision that is heatlhy D_0

→ True Positive Fraction

$Pr(P|W) = Pr(D_1|\theta_1)$
(sensitivity)

→ True negative

$Pr(P^c|W^c) = Pr(D_0|\theta_0)$
(specificity)

→ False Positive Fraction

$Pr(P|W^c) = Pr(D_1|\theta_0)$

→ False negative.

$Pr(P^c|W) = \cdots$

Facts: $\begin{cases} Pr(P^c|W^c) + Pr(P|W^c) = 1 \\ Pr(P|W^c) + Pr(P^c|W^c) = 1 \end{cases}$

$Pr(P^c|W) + Pr(P^c|W^c) \neq 1$

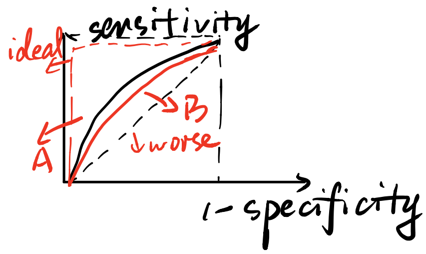
Positive Predictive value (PPV)

$Pr(W|P) \neq Pr(P|W)$
posterior prob

$$Pr(W|P) = \frac{Pr(P|W)\,Pr(W)}{Pr(P)} = \frac{Pr(P|W)\,Pr(W)}{Pr(P|W)\,P(W) + Pr(P|W^c)\,P(W^c)}$$

→ ROC analysis (Tradeoff)

(A more comprehensive characterization of binary statistical decision making



AUC: Area under Curve

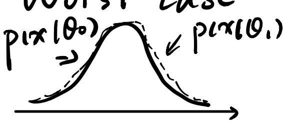
$p(D_1|\theta_0) = \int_{x_t}^{\infty} \underbrace{p(x|\theta_0)}_{} dx$ conditional distribution FP

$p(D_0|\theta_1) = \int_{-\infty}^{x_t} p(x|\theta_1) dx$ FN

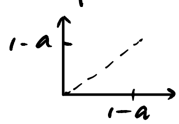
$p(D_1|\theta_1) = \int_{x_t}^{+\infty} p(x|\theta_1) dx$ TP

$p(D_0|\theta_0) = \int_{-\infty}^{x_t} p(x|\theta_0) dx$ TN

Worst case

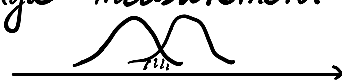


$Pr(x \le x_t|\theta_1) = a$ $Pr(x \ge x_t|\theta_0) = 1-a$

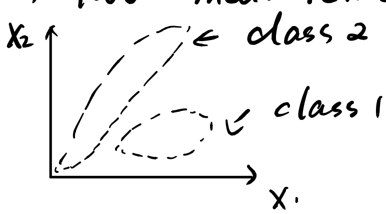


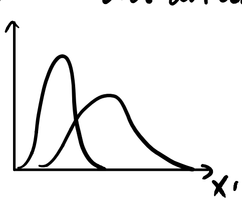
→ Binary decision

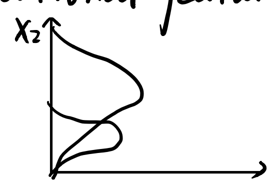
⇒ single measurement (single biomarker)



⇒ Two measurements (add an additional feature)







⇒ Multiple measurements. (more features)

$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \in R^N \xrightarrow{\phi} \phi(x) \in R^M$ $M >> N$ Dimensionality lifting

→ Discriminant Analysis (Optimal statistical decision making)

Problem: choosing the optimal threshold x_t

Definition: $\varepsilon \triangleq Pr(\theta_0) Pr(D_1|\theta_0) + Pr(\theta_1) Pr(D_0|\theta_1)$?

Bayesian risk:

① Cost C_{ij} the cost of making decision i given that the subject belongs to j

$C_{ij}\ (i \neq j) > C_{ij}\ (i = j)$

② Cost for binary decision $i, j \in \{0, 1\}$

$\begin{cases} C_{01},\ C_{10} = 1 \\ C_{00},\ C_{11} = 0 \end{cases}$ (for example) or $\begin{cases} C_{01}\ C_{10} = \inf \\ C_{00}\ C_{11} = 0 \text{ or } 1 \end{cases}$

Approach: Find an x_t that minimize the following "average" cost/risk (τ)

$$\tau = [C_{00} Pr(D_0|\theta_0) + C_{10} Pr(D_1|\theta_0)] Pr(\theta_0) + Pr(\theta_1) \cdot [C_{01} Pr(D_0|\theta_1) + C_{11} Pr(D_1|\theta_1)]$$

minimize w.r.t. x_t

$\Downarrow$ simplify $C_{00} = C_{11} = 0$

$$\tau = Pr(\theta_0) C_{10} Pr(D_1|\theta_0) + Pr(\theta_1) C_{01} Pr(D_0|\theta_1)$$

$$= Pr(\theta_0) C_{10} \underbrace{\int_{x_t}^{\infty} p(x|\theta_0)\, dx}_{\Omega_1:\ 1 - \int_{\Omega_0} p(x|\theta_0)\, dx} + Pr(\theta_1) C_{01} \underbrace{\int_{-\infty}^{x_t} p(x|\theta_1)\, dx}_{\Omega_0}$$

$$\tau = Pr(\theta_0) C_{10} + \underbrace{\int_{\Omega_0} [Pr(\theta_1) C_{01} p(x|\theta_1) - Pr(\theta_0) C_{10} p(x|\theta_0)]\, dx}_{\min\ (\Delta)}$$

for an x

a) $Pr(\theta_1) C_{01} p(x|\theta_1) < Pr(\theta_0) C_{10} p(x|\theta_0)$

if $x \in \Omega_0$, (Δ) will decrease.

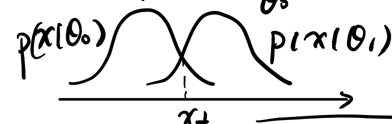
other wise

b) $Pr(\theta_1) C_{01} p(x|\theta_1) > Pr(\theta_0) C_{10} p(x|\theta_0)$

if $x \in \Omega_0$, (Δ) will increase

Thus, decision rule

$$Pr(\theta_1) C_{01} p(x|\theta_1) \underset{\theta_0}{\overset{\theta_1}{\gtrless}} Pr(\theta_0) C_{10} p(x|\theta_0)$$



For 0-1 cost & uniform prior, x_t is the optimal threshold.

⇒ likelihood ratio test

$$R = \frac{Pr(\theta_1) C_{01}\, p(x|\theta_1)}{Pr(\theta_0) C_{10}\, p(x|\theta_0)} \underset{\theta_0}{\overset{\theta_1}{\gtrless}} 1$$

if $Pr(\theta_1) = Pr(\theta_0)$ & $C_{01} = C_{10}$
$\Downarrow$
maximum likelihood test

→ Example

$x \sim N(\mu_{\theta_0}, K_{\theta_0})$. healthy

$x \sim N(\mu_{\theta_1}, K_{\theta_1})$ diseased

$$p(x|\theta_1) = (2\pi)^{-\frac{N}{2}} (\det K_{\theta_1})^{-\frac{1}{2}} \exp[-\tfrac{1}{2}(x-\mu_{\theta_1})^T K_{\theta_1}^{-1}(x-\mu_{\theta_1})]$$

$$\ln \frac{p(x|\theta_1)}{p(x|\theta_0)} = \ln \frac{(\quad)}{(\quad)} \gtrless \ln 1 \qquad \ln p(\vec{x}|\theta_1) - \ln p(\vec{x}|\theta_0) \gtrless 0$$

0
$\Uparrow$
0

$$\underline{\tfrac{1}{2}(x-\mu_{\theta_0})^T K_{\theta_0}^{-1}(x-\mu_{\theta_0}) - \tfrac{1}{2}(x-\mu_{\theta_1})^T K_{\theta_1}^{-1}(x-\mu_{\theta_1})} + \tfrac{1}{2}\ln\frac{\det K_{\theta_0}}{\det K_{\theta_1}} \gtrless 0$$

$$\Leftrightarrow (\mu_{\theta_1}-\mu_{\theta_0})^T K^{-1} x \gtrless \underbrace{\tfrac{1}{2}\mu_{\theta_1}^T K^{-1}\mu_{\theta_1} - \tfrac{1}{2}\mu_{\theta_0}^T K^{-1}\mu_{\theta_0}}_{\text{Constant}}$$

Futher simplification : $K = \sigma^2$ $\quad K^{-1} = \frac{1}{\sigma^2}$ $\quad \mu_{\theta_0} = \mu_0$ $\quad \mu_{\theta_1} = \mu_1$

→ signal-to-noise ratio (SNR)

$$\underset{\text{data collected}}{g'(t)} = \underset{\text{ture}}{g(t)} + \underset{\text{noise}}{e(t)} \qquad I'(x) = I(x) + \overset{\text{noise}}{e(x)}$$

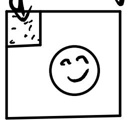
$$SNR \triangleq \begin{cases} \dfrac{\text{mean}(I(x))}{\sigma(e(x))} \text{ or } \dfrac{\text{mean}(g(t))}{\sigma(e(x))} \to \sigma \text{ (WSS)} \\[2ex] \dfrac{\max(I(x))}{\sigma} \\[2ex] \dfrac{\sigma(I(x))}{\sigma_{noise}} \quad I(x) \text{ is stochastic} \end{cases}$$

Estimation of SNR (noise i.i.d.)

① Estimation of mean

$\hat{I} = \frac{1}{N_\Omega} \sum_{i \in \Omega} I(x_i)$ $\Omega \triangleq$ region where the signal is homogenous

region of no signal



$s^2 = \frac{1}{N_\Omega - 1} \sum_{i \in \Omega} (e(x_i) - \bar{e})^2 = \sigma^2$

signal averging $I'_n(x) = I_n(x) + e_n(x)$ (n means the nth repeat)

averaging $\frac{1}{N} \sum_{n=1}^{N} I_n'(x) = \frac{1}{N} \sum \underbrace{I_n(x)}_{I(x)} + \frac{1}{N} \sum e_n(x)$

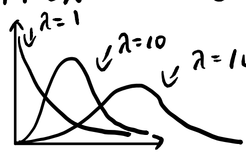
$Var[e_n(x)] = \sigma^2$ $Var[\frac{1}{N} \sum e_n(x)] = \frac{\sigma^2}{N} \xrightarrow{std} \frac{\sigma}{\sqrt{N}}$

Thus, N signal averages leads a factor of $\sqrt{N}$ SNR improvement.

Possion Random Variable

→ A discrete R.V. X

$Pr(X = n) = e^{\lambda} \frac{\lambda^n}{n!} = Pois(\lambda)$ PMF $E[X] = \lambda$ $Var[X] = \lambda$



$CF = E[e^{-i2\pi u x}] = \sum_{n=0}^{\infty} e^{-i2\pi u n} e^{\lambda} \frac{\lambda^n}{n!} = \exp[\lambda(e^{-i2\pi u} - 1)]$

Given X Y both Possion R.V. independent $Z = \underset{(\lambda_1)}{X} + \underset{(\lambda_2)}{Y}$

Q: What's the distribiution of Z

$C_Z(u) = C_X(u) C_Y(u) = \exp[\lambda_1 (e^{-i2\pi u} - 1)] \exp[\lambda_2 (e^{-i2\pi u} - 1)]$

$= \exp[(\lambda_1 + \lambda_2) e^{-i2\pi u} - (\lambda_1 + \lambda_2)]$

still Possion R.V. $\lambda = \lambda_1 + \lambda_2$

Modeling Biological System (Dynamic System Modeling

$f(t) \rightarrow$ Measurement System $\rightarrow g(t) \rightarrow \oplus \rightarrow g'(t) \rightarrow$ Recovery System $\rightarrow \hat{f}(t)$

$f(t)$: Object function; $\oplus$: $n(t)$ radon perturbation; $g'(t)$: Actual data

Object function $\leftarrow$ Biological System $\leftarrow$ Actual data

- Describe the dynamic behavior of system
- Determine some key properties of the system.
- Predict behavior

$\rightarrow$ Differential Equations (ODE & PDE)

$f(t)$: stimulus or source $\quad g(t)$: response / behavior of the system.

$$\frac{d^n g(t)}{dt^n} + k_1(t)\frac{d^{n-1}g(t)}{dt^{n-1}} + k_2(t)\frac{d^{n-2}g(t)}{dt^{n-2}} + k_n(t)g(t) = f(t)$$

$\Rightarrow$ ODE: ordinary differential equation

All functions & derivatives defined w.r.t. the same variable.

PDE: partial differential equation $u(x,t)$

$$\frac{\partial u(r,t)}{\partial(t)} = a\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) \quad \text{(Heat Equation)}$$

wave equation: $\dfrac{\partial^2 u(x,t)}{\partial t^2} = c^2\left(\dfrac{\partial^2 u}{\partial x^2}\right)$

ODE:

Remarks

① linear ODE: $\left(\dfrac{dg(t)}{dt}\right)^2 + kg^2(t) = cf(t)$ not linear ODE

② $f(t)$: source $\quad g(t)$: response

③ linear ODE $\Longleftrightarrow$ linear System

④ Define a operator $D = \left[\dfrac{d^n g(t)}{dt^n} + k\dfrac{d^{n-1}(t)}{dt^{n-1}} + \cdots + g(t)\right]$

$Dg(t) = f(t) \qquad \dfrac{d^n g(t)}{dt^n} + k_1\dfrac{d^{n-1}g(t)}{dt^{n-1}} \cdots + k_n = f(t) \in$ Delta function

$\vee g(t)$ for $\delta(t)$ as the input $\Rightarrow h(t)$

For any $g(t) = \int_{-\infty}^{+\infty} \underbrace{h(t-t')}_{\text{impulse response systems}} f(t')\,dt'$

$\rightarrow$ Example

W