Information Transfer

S. Ryffel

14th October 2006

Contents

1 Review of Probability Theory

1.1 Probability Models

Probability Model or probability experiment consists of:
- **Sample Space** Ω set of possible outcomes, all sample points ω
- **Event** E subset of the sample space, has a probability
- **Rule** assigns probabilities to events

Axioms of Probability are three restrictions:
1. probability of each event lies between 0 and 1
2. probability of the sample space is 1
3. $\Pr(\cup_i E_i) = \sum_i \Pr(E_i)$ union of any sequence of disjoint events E_i
   $\rightarrow \Pr(A^c) = 1 - \Pr(A)$
   $\rightarrow \Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(AB)$

Complementary Event E^c set of all sample points not in E

1.2 Conditional Probabilities

Conditional Probability of A, conditional on B

$$\Pr(A|B) = \frac{\Pr(AB)}{\Pr(B)}$$

$\rightarrow \Pr(A|B)\Pr(B) = \Pr(B|A)\Pr(A)$

Independent A and B are independent iff

$$\Pr(AB) = \Pr(A)\cdot\Pr(B) \iff \Pr(A|B) = \Pr(A) \Leftrightarrow \Pr(B|A) = \Pr(B)$$

- In general: independent $\Longrightarrow$ uncorrelated

Uncorrelated $\mathrm{Cov}[X,Y] = \mathrm{E}\left[(X-\overline{X})(Y-\overline{Y})\right] = 0$

Conditionally Independent given C

$$\Pr(AB|C) = \Pr(A|C)\Pr(B|C)$$

1.3 Random Variables

Random Variable X function $\omega \in \Omega \rightarrow \mathbb{R}$
- complex random variable: $\omega \in \Omega \rightarrow \mathbb{C}$
- vector random variable: $\omega \in \Omega \rightarrow \mathbb{R}^n$
- **Continuous Random Variable** has a finite density
- **Defective Random Variable** X there is a set of sampling points of positive probability for which the mapping is either undefined or defined to be $\pm\infty$.
- **Discrete Random Variable** has a finite number of possible outcomes. $\{P_X(x_i)\}$ is called the probability mass function of X.

Distribution Function $F_X(x)$ of the random variable X

$$F_X(x) = \Pr(X \le x) = \Pr(\{\omega \in \Omega \mid X(\omega) \le x\})$$

Events with zero probability have no effect on F_X.

Probability Density $f_X(x)$ of X

$$f_X(x) = \frac{\partial}{\partial x} F_X(x) \Longleftrightarrow F_X(x) = \int_0^x f_X(t)\,\mathrm{d}t$$

Joint Distribution Function of n random variables X_i

$$F_{X_1,\dots,X_n}(x_1,\dots,x_n) = \Pr(X_1 \le x_1,\dots,X_n \le x_n)$$

$$F(x_1,\dots,x_n) = \prod_{i=1}^{n} \Pr(X_i \le x_i) \iff X_1,\dots,X_n \text{ are independent}$$

Joint Probability Density

$$f(x_1,\dots,x_n) = \frac{\partial^n F(x_1,\dots,x_n)}{\partial x_1 \dots \partial x_n}$$

Joint PMF

$$\Pr(X_1 = x_1,\dots,X_n = x_n)$$

1.4 Expectations

Expected Value or the mean

$$\begin{aligned}\mathrm{E}[X] = \overline{X} &= \int_{-\infty}^{\infty} x \cdot f_X(x)\,\mathrm{d}x \\ &= \sum_x x \cdot \Pr(x) \\ &= \int_0^{\infty} [1 - F_X(x)]\,\mathrm{d}x \qquad \text{wenn } F_X(x) = 0 \text{ für } x < 0\end{aligned}$$

Variance σ_X^2

$$\mathrm{Var}(X) = \sigma_X^2 = \mathrm{E}\left[(X - \overline{X})^2\right] = \mathrm{E}\left[X^2\right] - \overline{X}^2$$

Standard Deviation σ_X square root of the variance

$$\begin{aligned}&S_n = X_1 + X_2 + \dots + X_n \\ &\mathrm{E}[S_n] = \mathrm{E}[X_1] + \mathrm{E}[X_2] + \dots + \mathrm{E}[X_n] && \text{whether or not } X_i \text{ are independent} \\ &\sigma_{S_n}^2 = \textstyle\sum_{i=1}^n \sigma_{X_i}^2 && \text{for independent } X_i \\ &\mathrm{E}\left[\textstyle\prod_{i=1}^n X_i\right] = \textstyle\prod_{i=1}^n \mathrm{E}[X_i] && \text{for independent } X_i\end{aligned}$$

1.5 Random Variables as Functions of Random Variables

$$\begin{aligned}&\mathbf{Y} = \mathbf{g}(\mathbf{X}) \qquad \text{for a monotonic } g \\ &\Rightarrow F_X(x) = F_Y(g(x)) \qquad F_Y(y) = F_X(g^{-1}(y)) \\ &\Rightarrow f_X(x) = g'(x) f_Y(g(x)) \qquad f_Y(y) = \tfrac{f_X(g^{-1}(y))}{g'(g^{-1}(y))} \\ &\Rightarrow f_{\mathbf{Y}}(\mathbf{y}) = \frac{f_{\mathbf{X}}(g^{-1}(\mathbf{y}))}{\det\left(J_g(g^{-1}(\mathbf{y}))\right)} = \det(J_{g^{-1}}(\mathbf{y})) \cdot f_{\mathbf{X}}(g^{-1}(\mathbf{y}))\end{aligned}$$

$$\text{with } J_g(\mathbf{x}) = \begin{pmatrix} \frac{\partial y_1}{\partial x_1}(\mathbf{x}) & \dots & \frac{\partial y_1}{\partial x_n}(\mathbf{x}) \\ \vdots & & \vdots \\ \frac{\partial y_n}{\partial x_1}(\mathbf{x}) & \dots & \frac{\partial y_n}{\partial x_n}(\mathbf{x}) \end{pmatrix}$$

$$\begin{aligned}&\mathbf{Z} = \mathbf{X} + \mathbf{Y} \qquad \textit{for } X \text{ and } Y \text{ independent} \\ &\Rightarrow F_Z(z) = \int_{-\infty}^{\infty} F_X(z-y)\,\mathrm{d}F_Y(y) = \int_{-\infty}^{\infty} F_Y(z-x)\,\mathrm{d}F_X(x) \\ &\Rightarrow f_Z(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)\,\mathrm{d}y = \int_{-\infty}^{\infty} f_Y(z-x) f_X(x)\,\mathrm{d}x\end{aligned}$$

1.6 Transforms

Moment Generating Function (MGF) of a random variable X

$$g_X(r) = \mathrm{E}\left[e^{rX}\right] = \int_{-\infty}^{\infty} e^{rx}\,\mathrm{d}F_X(x)$$

If $g_X(r)$ exists in a region of real r around 0, then:

$$\frac{\partial^n g_X(r)}{\partial r^n} = \int_{-\infty}^{\infty} x^n e^{rx}\,\mathrm{d}F_X(x)\,; \qquad \left.\frac{\partial^n g_X(r)}{\partial r^n}\right|_{r=0} = \mathrm{E}[X^n]$$

$S = X_1 + X_2 + \dots + X_n$ is a sum of independent random variables, then:

$$g_S(r) = \mathrm{E}\left[e^{rS}\right] = \prod_{i=1}^{n} g_{X_i}(r)$$

Characteristic Function of a random variable X is the inverse Fourier transform of the density function

$$\mathrm{E}\left[e^{j\omega}\right] = \int_{-\infty}^{\infty} e^{j\omega}\,\mathrm{d}F_X(x)$$

1.7 Basic Inequalities

Markov inequality (is typically very weak)

$$\Pr(Y \ge y) \le \frac{\mathrm{E}[Y]}{y}$$

Chebyshev inequality

$$\Pr(|Z - \mathrm{E}[Z]| \ge \varepsilon) = \frac{\sigma_Z^2}{\varepsilon^2}$$

Exponential bound or Chernoff bound. Can be optimized over r to get the strongest bounds

$$\begin{aligned}&\Pr(Z \ge a) \le g_Z(r)\exp(-ra) && \text{for } r \ge 0 \\ &\Pr(Z \le a) \le g_Z(r)\exp(-ra) && \text{for } r \le 0\end{aligned}$$

1.8 Complex Random Variables

Expectation $\mathrm{E}[C] = \mathrm{E}[\Re(C)] + i\mathrm{E}[\Im(C)]$ and $\mathrm{E}\left[|C|^2\right] = \mathrm{E}[(\Re(C))^2] + \mathrm{E}[(\Im(C))^2]$

Variance $\mathrm{Var}(C) = \mathrm{E}\left[|C - \mathrm{E}[C]|^2\right] = \mathrm{E}\left[|C|^2\right] - |\mathrm{E}[C]|^2$

1.9 Table of Standard Random Variables

Name	Density	Mean	Variance	MGF
Exponential	$\lambda \exp(-\lambda x);\quad x \geq 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda - r}$
Rayleigh	$\frac{x}{\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2}\right)$	$\sigma\frac{\pi}{2}$	$\frac{4-\pi}{2}\sigma^2$	
Erlang	$\frac{\lambda^n x^{n-1}\exp(-\lambda x)}{(n-1)!};\quad x \geq 0$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$	$\left(\frac{\lambda}{\lambda - r}\right)^n$
Gaussian	$\frac{1}{\sigma\sqrt{2\pi}}\exp\left(\frac{-(x-a)^2}{2\sigma^2}\right)$	a	σ^2	$\exp(ra + r^2\sigma^2/2)$
Uniform	$\frac{1}{a};\quad 0 \leq x \leq a$	$\frac{a}{2}$	$\frac{a^2}{12}$	$\frac{\exp(ra)-1}{ra}$

Name	PMF	Mean	Variance	MGF
Bernoulli	$\Pr_N(0) = 1 - p;\quad \Pr_N(1) = p$	p	$p(1-p)$	$1 - p + pe^r$
Binomial	$\frac{k!}{n!(k-n)!}p^n(1-p)^{k-n};\quad 0 \leq k \leq n$	kp	$kp(1-p)$	$(1 - p + pe^r)^k$
Geometric	$(1-p)p^n;\quad n \geq 0$	$\frac{p}{1-p}$	$\frac{p}{(1-p)^2}$	$\frac{1-p}{1-pe^r}$
Poisson	$\frac{\lambda^n \exp(-\lambda)}{n!};\quad n \geq 0$	λ	λ	$\exp[\lambda(e^r - 1)]$

Table 1: Table of Standard Random Variables

2 The Inner Product – a First Glimpse

2.1 The Inner Product

$$\begin{aligned} \langle g, h\rangle &= \int_{-\infty}^{\infty} g(t)h^*(t)\,\mathrm{d}t && \text{inner Product} \\ &= \int_{-\infty}^{\infty} \Re\left(g(t)h^*(t)\right)\mathrm{d}t + i\int_{-\infty}^{\infty} \Im\left(g(t)h^*(t)\right)\mathrm{d}t \\ \|g\|^2 &= \langle g, g\rangle = \int_{-\infty}^{\infty} |g(t)|^2\,\mathrm{d}t && \text{energy of a complex signal} \end{aligned}$$

Properties of the inner product, valid if all functions are of finite energy.

$$\begin{aligned} \langle g, h\rangle &= \langle h, g\rangle^* \\ \langle \alpha g, h\rangle &= \alpha\langle g, h\rangle \\ \langle g, \alpha h\rangle &= \alpha^*\langle g, h\rangle \\ \langle g_1 + g_2, h\rangle &= \langle g_1, h\rangle + \langle g_2, h\rangle \end{aligned}$$

2.2 Integrating a Complex Signal

$$\begin{aligned} \int g(t)\,\mathrm{d}t &= \int \Re(g(t))\,\mathrm{d}t + i\int \Im(g(t))\,\mathrm{d}t \\ \int g^*(t)\,\mathrm{d}t &= \left(\int g(t)\,\mathrm{d}t\right)^* \\ \Re\left(\int g(t)\,\mathrm{d}t\right) &= \int \Re(g(t))\,\mathrm{d}t \\ \Im\left(\int g(t)\,\mathrm{d}t\right) &= \int \Im(g(t))\,\mathrm{d}t \\ \int \left|g(t)h(t)\right|\,\mathrm{d}t &= \int \left|g(t)\right| \cdot |h(t)|\,\mathrm{d}t \\ \left|\int g(t)h(t)\,\mathrm{d}t\right| &\leq \int \left|g(t)h(t)\right|\,\mathrm{d}t \end{aligned}$$

Integrable $g(\cdot)$ is called integrable iff

$$\int_{-\infty}^{\infty} \left|g(t)\right|\,\mathrm{d}t \leq \infty$$

2.3 The Cauchy-Schwarz Inequality

For $\|g\|, \|h\| < \infty$:

$$\left|\langle g, h\rangle\right| \leq \|g\| \cdot \|h\|$$

Equality holds for two functions that are $\mathcal{L}_2$ indistinguishable from a scaled version of each other: $g = \alpha \cdot h$

3 Convolutions and Filters

3.1 The Definition of the Convolution

$$(x \star h)(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)\,\mathrm{d}\tau$$

If x and h are both of finite energy and if one of them is also integrable, then $x \star h$ is a finite energy signal.

Properties of the Convolution

$$\begin{aligned} x \star h &= h \star x && x, h \in \mathcal{L}_1(\mathbb{R}) \\ (x \star g) \star h &= x \star (g \star h) && x, g, h \in \mathcal{L}_1(\mathbb{R}) \\ x \star (g + h) &= x \star g + x \star h && x, g, h \in \mathcal{L}_1(\mathbb{R}) \\ x \star (\alpha g + \beta h) &= \alpha(x \star g) + \beta(x \star h) && x, g, h \in \mathcal{L}_1(\mathbb{R}) \end{aligned}$$

Graphic Convolution of $g(t) = f \star h$

1. Mirror $h(t) \to h(-t)$
2. Shift $h(-t)$ from $t = -\infty$ to $t = \infty$
3. The value of $(f \star h)$ at position t is the multiplication of the areas of $f(\cdot)$ and $h(\cdot - t)$.

3.2 Filter

Stable Filter (BIBO) The impulse response h of the filter is integrable.
Causal Filter $h(t) = 0$ for $t < 0$.

3.3 The Matched Filter

Matched Filter for the complex waveform $g(\cdot)$ is

$$\check{g}^*(t) = g^*(-t)$$

Inner Product via Matched Filtering between the energy limited functions u and v.

$$\langle u, v\rangle = (u \star \check{v})(0)$$

$$\int_{-\infty}^{\infty} u(t)v^*(t-t_0)\,\mathrm{d}t = (u \star \check{v})(t_0)$$

3.4 The Unit-Gain Ideal Low-Pass Filter

$$\begin{aligned}
\mathrm{LPF}_{W_c}(t) &= \begin{cases} 2W_c \frac{\sin(2\pi W_c t)}{2\pi W_c t} & \text{if } t \neq 0 \\ 2W_c & \text{if } t = 0 \end{cases} \\
&= 2W_c \cdot \mathrm{sinc}(2W_c t) \\
\mathrm{sinc}(\xi) &= \begin{cases} \frac{\sin(\pi\xi)}{\pi\xi} & \text{if } \xi \neq 0 \\ 1 & \text{if } \xi = 0 \end{cases} \qquad \xi \in \mathbb{R} \\
\widehat{\mathrm{LPF}}_{W_c}(f) &= \mathrm{I}\left\{|f| \leq W\right\}
\end{aligned}$$

$$x \star \mathrm{LPF}_W \star \mathrm{LPF}_W = x \star \mathrm{LPF}_W$$

$$x \star \mathrm{LPF}_W = \int_{-W}^{W} \hat{x}(f)e^{i2\pi ft}\,\mathrm{d}f$$

- The impulse response is not an integrable function and thus not stable.
- The impulse response is however of finite energy.

3.5 The Unit-Gain Ideal Band-Pass Filter

Of bandwidth W around the carrier frequency f_c, where $f_c > W/2 > 0$.

$$\begin{aligned}
\mathrm{BPF}_{W,f_c}(t) &= 2W \cdot \cos(2\pi f_c t) \cdot \mathrm{sinc}(Wt) \\
&= 2\Re\left(\mathrm{LPF}_W(t)e^{i2\pi f_c t}\right) \\
\widehat{\mathrm{BPF}}_{W,f_c}(f) &= \mathrm{I}\left\{\left||f| - f_c\right| \leq W/2\right\}
\end{aligned}$$

This filter is too non-stable and non-causal.

3.6 Mathematical Notes on $\mathcal{L}_p(\mathbb{R})$

$g(t)$ is in $\mathcal{L}_p(\mathbb{R})$, iff

$$\int_{-\infty}^{\infty} |g(t)|^p\,\mathrm{d}t < \infty \qquad \Rightarrow \qquad \|g(t)\|_p = \left(\int_{-\infty}^{\infty} |g(t)|^p\,\mathrm{d}t\right)^{1/p}$$

$g(t)$ is in $\mathcal{L}_\infty(\mathbb{R})$, iff there exists a constant $\mathcal{M}$ such that $|g(t)| \leq \mathcal{M} \quad \forall t \in \mathbb{R}$.
For $h \in \mathcal{L}_1(\mathbb{R})$ and $x \in \mathcal{L}_p(\mathbb{R})$:

$$\|h \star x\|_p \leq \|h\|_1 \cdot \|g\|_p$$

4 The Frequency Response of Filters and Bandlimited Signals

4.1 Review of the Fourier Transform

Fourier Transform

$$\hat{x}(f) = \int_{-\infty}^{\infty} x(t)e^{-i2\pi ft}\,\mathrm{d}t$$

Inverse Fourier Transform

$$x(t) = \int_{-\infty}^{\infty} \hat{x}(f)e^{i2\pi ft}\,\mathrm{d}f$$

Parseval's Theorem If the signals x and y are of finite energy, then

$$\begin{aligned}
\langle x, y\rangle &= \langle \hat{x}, \hat{y}\rangle \\
\|x\| &= \|\hat{x}\| \\
\int_{-\infty}^{\infty} |x(t)|^2\,\mathrm{d}t &= \int_{-\infty}^{\infty} |\hat{x}(t)|^2\,\mathrm{d}t
\end{aligned}$$

Properties of the Fourier Transform (see table 2)

- Symmetries:

$$\begin{aligned}
\hat{x}(-f) = \hat{x}^*(f) \quad &\Longleftrightarrow \quad x \text{ is a real signal} \\
\hat{x} \text{ is conjugate symmetric} \quad & \\
\hat{x}(-f) = -\hat{x}^*(f) \quad &\Longleftrightarrow \quad x \text{ is a imaginary signal} \\
\hat{x} \text{ is conjugate anti-symmetric} \quad &
\end{aligned}$$

- $\hat{\hat{x}} = \check{x}$ Therefore if $\xi(-t) = -\xi(t)\ \forall t$ and $\hat{\xi}(\cdot) = \eta(\cdot)$, then $\hat{\eta}(\cdot) = \xi(\cdot)$.
- If two integrable functions have the *same Fourier Transform*, then they are *indistinguishable*.

Property	x	$\hat{x}$
linearity	$\alpha x + \beta y$	$\alpha\hat{x} + \beta\hat{y}$
time shifting	$x(t - t_0)$	$e^{i2\pi f t_0}\hat{x}(f)$
frequency shifting	$e^{i2\pi f_0 t}x(t)$	$\hat{x}(f - f_0)$
conjugation	$x^*(t)$	$\hat{x}^*(-f)$
scaling	$x(\alpha t)$	$\frac{1}{\lvert\alpha\rvert}\hat{x}\left(\frac{f}{\alpha}\right)$
convolution in time	$x \star y$	$\hat{x}(f)\hat{y}(f)$
multiplication in time	$x(t)y(t)$	$\hat{x} \star \hat{y}$
real part	$\Re(x(t))$	$\frac{1}{2}\hat{x}(f) + \frac{1}{2}\hat{x}^*(f)$
	$\frac{1}{2}\left(x(t) + x^*(-t)\right)$	$\Re(\hat{x}(f)$
imaginary part	$\Im(x(t))$	$\frac{1}{2}i\hat{x}(f) + \frac{1}{2}\hat{x}^*(-f)$
	$\frac{1}{2}\left(x(t) - x^*(-t)\right)$	$i\Im(\hat{x}(f))$
derivative in time	$\frac{\mathrm{d}^n x(t)}{\mathrm{d}t^n}$	$(2\pi i f)^n\hat{x}(f)$
multiplication by t	$tx(t)$	$i\frac{1}{2\pi}\frac{\mathrm{d}}{\mathrm{d}f}\hat{x}(f)$
cosine	$\cos(2\pi f_0 t)$	$\frac{1}{2}\left(\delta(f + f_0) + \delta(f - f_0)\right)$
sine	$\sin(2\pi f_0 t)$	$\frac{1}{2}i\left(\delta(f + f_0) - \delta(f - f_0)\right)$

Table 2: Properties of the Fourier Transform

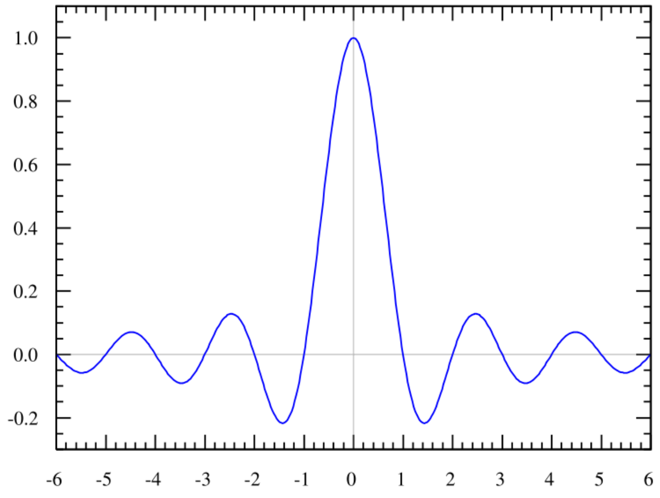

Figure 1: Normalized $\mathrm{sinc}(\xi) = \frac{\sin(\xi\pi)}{\xi\pi}$

4.2 On Brickwalls and Sincs

$$\delta \cdot \mathrm{sinc}(\alpha t) \quad \text{and} \quad \beta \cdot \mathrm{I}\left\{\lvert f\rvert \le \gamma\right\}$$

are Fourier pairs, iff

$$\begin{aligned} \delta &= 2\gamma\beta \\ \alpha &= 2\gamma \\ &\text{or} \\ \beta &= \tfrac{\delta}{\alpha} \\ \gamma &= \tfrac{\alpha}{2} \end{aligned}$$

One can derive it from the Inverse Fourier Transform of the Brickwall function.

4.3 The Frequency Response of a Filter

x and h	$\to x \star h$
both integrable	$\to$ integrable
one integrable	$\to$ of finite energy
other of finite energy	

$$\widehat{x \star y}(f) = \hat{x}(f) \cdot \hat{y}(f)$$

Frequency Response is the Fourier Transform of the Impulse Response. If the Impulse Response is either integrable or of finite energy.

4.4 Low Pass Filtering and Bandlimited Signals

4.4.1 Energy-Limited Signals

Energy Limited Bandlimited Signals $x(\cdot) \in \mathcal{L}_2(\mathbb{R})$ is bandlimited to W Hz iff

(a) "'$x(\cdot)$ is unaltered by the LPF_w:'"

$$(x \star \mathrm{LPF}_W)(t) = x(t), \quad t \in \mathbb{R} \quad \Longleftrightarrow \quad x(t) = \int_{-W}^{W} \hat{x}(f)e^{i2\pi f t}\,\mathrm{d}f$$

(b) *Equivalent definition for bandlimited functions in $\mathcal{L}_2(\mathbb{R})$:*
If $x \in \mathcal{L}_2(\mathbb{R})$ is bandlimited to W Hz then it can be expressed in the form

$$x(t) = \int_{-W}^{W} g(f)e^{i2\pi f t}\,\mathrm{d}f$$

where g satisfies

$$\int_{-W}^{W} \lvert g(f)\rvert^2\,\mathrm{d}f < \infty$$

and

$$g(\cdot) = \hat{x}(\cdot)$$

Bandwidth in the baseband of an energy-limited signal x is the smallest frequency W to which x is bandlimited.

If $x \in \mathcal{L}_2(\mathbb{R})$ is bandlimited to W Hz then x is a continuous function and all it's energy is contained in the frequency interval:

$$\int_{-\infty}^{\infty} \lvert\hat{x}(f)\rvert^2\,\mathrm{d}f = \int_{-W}^{W} \lvert\hat{x}(f)\rvert^2\,\mathrm{d}f$$

$y = x \star LPF_W$:

1. $\|y\|_2 \leq \|x\|_2$
2. $\hat{y}$ is indistinguishable from $\hat{x} \cdot \mathrm{I}\{|f| \leq W\}$
3. All of y energy is contained in the frequency interval.
4. The Inverse Fourier Transform formula is valid for y at every time instance.

4.5 The Bandwidth of a Product of two Signals

Let the finite energy signals x_1 and x_2 be bandlimited to W_1 Hz and W_2 Hz respectively. Then the product signal $x_1(t)x_2(t)$ is of finite energy and bandlimited to $W_1 + W_2$ Hz.

5 Passband Signals and their Representation

5.1 The Passband Signal

Passband Signal A signal x in $\mathcal{L}_2(\mathbb{R})$ or in $\mathcal{L}_1(\mathbb{R})$ is bandlimited to W Hz around the carrier frequency $f_c > \frac{W}{2} > 0$

$$x(t) = (x \star \mathrm{BPF}_{w,f_c})(t), \qquad t \in \mathbb{R}$$
$$\iff \qquad \hat{x}(f) = 0, \quad \forall \ \big||f| - f_c\big| \leq W/2$$

Multiplication by a carrier doubles the bandwidth Given a baseband signal x with bandwidth W Hz

$$y(t) = x(t)\cos(2\pi f_c t) \quad \Longrightarrow \quad \hat{y}(f) = \frac{1}{2}\Big(\hat{x}(f - f_c) + \hat{x}(f + f_c)\Big)$$

y has bandwidth $2W$.

5.2 The Analytic Signal

$$\hat{x}_A(f) = \begin{cases} \hat{x}_{PB}(f) & \text{for } f \geq 0 \\ 0 & \text{for } f < 0 \end{cases}$$

Properties

- $\|x_A\|^2 = \frac{1}{2} \ \|x_{PB}\|^2$
- $\hat{x}_{PB}(f) = \hat{x}_A(f) + \hat{x}_A^*(-f)$
  $x_{PB} = 2\Re(x_A(t))$
- We look at real passband signals only. x_A is complex.

5.3 The Baseband Representation

Baseband Representation of a real x_{PB} which is bandlimited to W Hz around f_c

$$\begin{aligned} x_{BB} &= \left(e^{-i2\pi f_c t} x_{PB}(t)\right) \star \mathrm{LPF}_{W_c} \\ &= x_I(t) + i\, x_Q(t) \\ x_I(t) &= \Re(x_{BB}(t)) = \left(x_{PB}(t)\cos(2\pi f_c t)\right) \star \mathrm{LPF}_{W_c} && \text{In-Phase Component} \\ x_Q(t) &= \Im(x_{BB}(t)) = -\left(x_{PB}(t)\sin(2\pi f_c t)\right) \star \mathrm{LPF}_{W_c} && \text{Quadrature Component} \end{aligned}$$

x_{BB} is bandlimited to $W/2$ Hz.

Properties of x_{BB}

- $\|x_{BB}\|^2 = \|x_A\|^2 = \frac{1}{2} \ \|x_{PB}\|^2$
- $\hat{x}_{PB}(f) = \hat{x}_{BB}(f - f_c) + \hat{x}_{BB}^*(-f - f_c)$
- time-domain relationship

$$\begin{aligned} x_{PB}(t) &= 2\Re\left(x_{BB}(t)e^{i2\pi f_c t}\right) \\ &= 2\Re(x_A(t)) \\ &= x_i(t)\cos(2\pi f_c t) - x_Q(t)\sin(2\pi f_c t) \end{aligned}$$

- inner products

$$\langle x_{PB}, y_{PB}\rangle = 2\Re\Big(\langle x_{BB}, y_{BB}\rangle\Big) = 2\Re\Big(\langle \hat{x}_{BB}, \hat{y}_{BB}\rangle\Big)$$

- Filtering $x \in \mathcal{L}_2(\mathbb{R})$ (W, f_c) with an integrable filter h.

$$\begin{aligned} (x \star h)_{BB} &= x_{BB} \star h'_{BB} \\ h' &= h \star \mathrm{BPF}_{W,f_c} \\ h'_{BB}(t) &= \left(e^{-i2\pi f_c t} h'(t)\right) \star \mathrm{LPF}_{W/2} \\ \widehat{(x \star h)}_{BB}(f) &= \hat{x}_{BB}(f) \cdot \hat{h}'_{BB}(f) \\ \hat{h}'_{BB}(f) &= \hat{h}(f + f_c) \cdot \mathrm{I}\{|f| \leq W/2\} \end{aligned}$$

- Low Pass Filter

$$\mathrm{LPF}_W \star \mathrm{LPF}_W = \mathrm{LPF}_W$$

- The cutoff frequency W_c of the LPF has to satisfy $W/2 \leq W_c \leq 2f_c - W/2$

5.4 Miscellaneous Results

- If the Fourier Transform of $x(\cdot)$ is zero for $|f| > W$

$$\begin{aligned} \int_{-\infty}^{\infty} x(t)e^{-i2\pi f t}\, \mathrm{d}t &= 0, && |f| > W \\ \int_{-\infty}^{\infty} x(t)\cos(2\pi f_c t)\, \mathrm{d}t &= 0, && |f_c| > W \\ \int_{-\infty}^{\infty} x(t)\sin(2\pi f_c t)\, \mathrm{d}t &= 0, && |f_c| > W \end{aligned}$$

- If the Fourier Transform of $\xi(\cdot)$ is zero for $|f| > W$, and $f_c > W/2$

$$\int_{-\infty}^{\infty} \xi(t)\cos^2(2\pi f_c t)\, \mathrm{d}t = \frac{1}{2}\int_{-\infty}^{\infty} \xi(t)\, \mathrm{d}t$$
$$\int_{-\infty}^{\infty} \xi(t)\cos(2\pi f_c t)\sin(2\pi f_c t)\, \mathrm{d}t = 0$$

6 Energy Limited Signals and the Sampling Theorem

6.1 Energy Limited Signals

$$\|u\|^2 = \int_{-\infty}^{\infty} |u(t)|^2\, \mathrm{d}t \qquad \text{energy in } u$$

Triangle Inequality

$$\|u+v\| \leq \|u\| + \|v\|$$

Equality holds for orthogonal signals: $\langle u, v\rangle = 0$.

6.2 The Geometry of $\mathcal{L}_2(\mathbb{R})$

Orthogonality u and v are orthogonal, if

$$\int_{-\infty}^{\infty} u(t)v^*(t)\,\mathrm{d}t = 0$$

Inner Product of u and v

$$\langle u, h\rangle = \int_{-\infty}^{\infty} u(t)v^*(t)\,\mathrm{d}t$$

Properties:

- $\langle u, v\rangle = \langle v, u\rangle^*$
- $\langle \alpha u + \beta v, w\rangle = \alpha\langle u, w\rangle + \beta\langle v, w\rangle$
- $\langle u, \alpha v + \beta w\rangle = \alpha^*\langle u, v\rangle + \beta^*\langle u, w\rangle$
- $\|u\|^2 = \langle u, u\rangle$
- $|\langle v, u\rangle| \leq \|v\| + \|u\|$

Projection the function w is the projection of $v \in \mathcal{L}_2(\mathbb{R})$ onto the function $u \in \mathcal{L}_2(\mathbb{R})$ if:

(a) $w \equiv \alpha u$ for some $\alpha \in \mathbb{C}$

(b) $v - w$ is orthonormal to u

$$\|v - w\| = \min_{\beta\in\mathbb{C}} \left\|v - \beta u\right\| = 0$$

$$\alpha = \frac{\langle v, u\rangle}{\|u\|^2}$$

6.3 Finite-Dimensional Linear Sub-Spaces

6.3.1 Linear Sub-Spaces

Linear Sub-Space consists of set of functions in $\mathcal{L}_2(\mathbb{R})$ with:

(a) $u \in \mathcal{U} \quad \Rightarrow \quad \alpha u \in \mathcal{U} \qquad \forall \alpha \in C$

especially $0 \in \mathcal{U}$ and $-u \in \mathcal{U}$

(b) $u, v \in \mathcal{U} \quad \Rightarrow \quad u + v \in \mathcal{U}$

Span $u_1, \ldots, u_n \in \mathcal{U}$ of $\mathcal{U}$

$$u \equiv \sum_{l=1}^{n} \alpha_l u_l \qquad \forall u \in \mathcal{U}$$

$$\mathcal{U} = \operatorname{span}\{u_1, \ldots, u_n\}$$

Linearly Independent $u_1, \ldots, u_n$ are linearly independent, if either holds:

(a) None of the functions can be written as a linear combination of the others:

$$u_l \equiv \sum_{l'\neq l} \beta_{l'} u_{l'}$$

(b)

$$\sum_{l=1}^{n} \alpha_l u_l \equiv 0 \quad \Longleftrightarrow \quad \alpha_i = 0$$

Basis $u_1, \ldots, u_n$ of a linear sub-space $\mathcal{U}$. u_i are linearly independent and span $\mathcal{U}$.
Dimension of a subspace $\mathcal{U}$ is d, if there exist d functions that form a basis for $\mathcal{U}$.

6.3.2 Orthonormal Bases

Definitions

Orthonormal A set of n functions $\phi_1, \ldots, \phi_n$ is said to be orthonormal, if

$$\left\langle \phi_l, \phi_{l'}\right\rangle = \begin{cases} 0 & \text{if } l \neq l' \\ 1 & \text{if } l = l' \end{cases}$$

orthonormal $\Rightarrow$ linearly independent

Orthonormal Basis $\phi_1, \ldots, \phi_n$ is a set of orthonormal functions, which spans a sub-space $\mathcal{U}$.

$$\begin{aligned} u &\equiv \textstyle\sum_{l=1}^{d} \underbrace{\left\langle u, \phi_l\right\rangle}_{\alpha_l} \phi_l && u \in \mathcal{U} \\ \|u\|^2 &= \textstyle\sum_{l=1}^{d} \left|\left\langle u, \phi_l\right\rangle\right|^2 && \\ \|v\|^2 &\geq \textstyle\sum_{l=1}^{d} \left|\left\langle v, \phi_l\right\rangle\right|^2 && v \in \mathcal{L}_2(\mathbb{R}) \\ \langle v, u\rangle &= \textstyle\sum_{l=1}^{d} \left\langle v, \phi_l\right\rangle \left\langle u, \phi_l\right\rangle^* && u, v \in \mathcal{U} \end{aligned}$$

Gram-Schmidt Procedure is used to produce an orthonormal basis $\phi_1, \ldots, \phi_n$ of $\mathcal{U}$ from any basis $u_1, \ldots, u_n$ with

$$\operatorname{span}\{u_1, \ldots, u_n\} = \operatorname{span}\{\phi_1, \ldots, \phi_n\} = \mathcal{U}$$

1. $\phi_1 = \frac{u_1}{\|u_1\|}$
2. $u_{m+1}^{\perp} = u_{m+1} - \sum_{l=1}^{m} \left\langle u_{m+1}, \phi_l\right\rangle \phi_l$
3. $\phi_{m+1} = \frac{u_{m+1}^{\perp}}{\|u_{m+1}^{\perp}\|}$

6.3.3 Projections

Definition w is the projection of v onto $\mathcal{U}$ if

(a) $w \in \mathcal{U}$ and

(b) $\langle v - w, u\rangle = 0, \quad u \in \mathcal{U}$

$$w \equiv \sum_{l=1}^{d} \left\langle v, \phi_l\right\rangle \phi_l$$

6.4 Infinite-Dimensional Linear Sub-Spaces

6.4.1 Complete Orthonormal Subspace

Complete or Closed The space $\mathcal{L}_2(\mathbb{R})$ is complete in the following sense: Every Cauchy Sequence in $\mathcal{L}_2(\mathbb{R})$ converges to some function in $\mathcal{L}_2(\mathbb{R})$. For closed sub-spaces, projection is well defined. (See Script p. 111)

Complete Orthonormal System (CONS) An orthonormal sequence of functions $\{\phi_l\}$ in $\mathcal{U}$ is said to be a **CONS for** $\mathcal{U}$ if the following equivalent conditions hold:

(a) *Projection*: Every $u \in \mathcal{U}$ can be arbitrarily well *approximated*. For every $\varepsilon > 0 \; \exists \; n_o$ with

$$\left\| u - \sum_{l=1}^{n} \alpha_l \phi_l \right\|^2 < \varepsilon \qquad \forall u \in \mathcal{U}, n > n_o, \quad \alpha_l = \left\langle u, \phi_l \right\rangle$$

(b) *Inner Product* for any $v \in \mathcal{L}_2(\mathbb{R})$ and $u \in \mathcal{U}$

$$\langle u, v \rangle = \lim_{n\to\infty} \sum_{l=1}^{n} \left\langle u, \phi_l \right\rangle \left\langle v, \phi_l \right\rangle^*$$

(c) The *Norm* of any $u \in \mathcal{U}$

$$\|u\|^2 = \lim_{n\to\infty} \sum_{l=1}^{n} \left| \left\langle u, \phi_l \right\rangle \right|^2$$

6.4.2 Examples

- The set of all $\mathcal{L}_2(\mathbb{R})$ functions that have zero energy outside some interval $[-\alpha, \alpha]$

$$\mathcal{L}_2([-\alpha,\alpha]) \triangleq \left\{ u : \int_{-\infty}^{\infty} |u(t)|^2 \, \mathrm{d}t = \int_{-\alpha}^{\alpha} |u(t)|^2 \, \mathrm{d}t \right\}$$

- *Functions of compact support*

$$\mathcal{L}_{cs}(\mathbb{R}) \triangleq \left\{ u \in \mathcal{L}_2(\mathbb{R}) : \int_{-\infty}^{\infty} |u(t)|^2 \, \mathrm{d}t = \int_{-\alpha(u)}^{\alpha(u)} |u(t)|^2 \, \mathrm{d}t \right\}$$

- The set of all functions of the form $p(t)e^{-|t|}$, where $p(t)$ is any polynomial.
- *The Fourier Series*: A CONS for all energy limited functions, that vanish outside the interval $[-W, W)$

$$\phi_l(\alpha) = \frac{1}{\sqrt{2W}} e^{i\pi l \alpha / W} \, \mathrm{I}\{-W \le \alpha < W\}$$

6.4.3 Projections onto Infinite-Dimensional Sub-Spaces

Let $\mathcal{U}$ be a closed linear sub-space of $\mathcal{L}_2(\mathbb{R})$, and let $v \in \mathcal{L}_2(\mathbb{R})$ be arbitrary. There exists some function $w \in \mathcal{U}$, called the projection of v onto $\mathcal{U}$ such that

(a) $v - w$ is orthogonal to any function in $\mathcal{U}$

(b) $\|v - w\| \le \|v - u\|, \quad \forall u \in \mathcal{U}$

$$w = \sum_{l=0}^{\infty} \left\langle v, \phi_l \right\rangle \phi_l$$

6.4.4 The Sampling Theorem

CONS for to W HZ bandlimited functions:

$$\hat{\phi}_l = \sqrt{T} e^{i2\pi l f T} \star \mathrm{LPF}_W \qquad T = \frac{1}{2W}$$

$$\phi_l = \frac{1}{\sqrt{T}} \mathrm{sinc}\left(\frac{t}{T} + l \right)$$

Let u be a energy-limited signal of bandwidth W Hz. Then

(a) u can be reconstructed from samples taken $T = \frac{1}{2W}$ seconds apart.

$$\lim_{n\to\infty} \left\| u - \sum_{l=-n}^{n} u(-lT) \mathrm{sinc}\left(\frac{t}{T} + l \right) \right\| = 0$$

(b) The energy of u is

$$\|u\|^2 = T \sum_{l=-\infty}^{\infty} |u(lT)|^2$$

(c) If v is an arbitrary energy-limited signal, and $\tilde{v}$ is the result of passing v through LPF_W

$$\langle v, u \rangle = \int_{-\infty}^{\infty} v(t) u^*(t) \, \mathrm{d}t = T \sum_{l=-\infty}^{\infty} \tilde{v}(lT) u^*(lT)$$

7 Sampling Passband Signals

Instead of sampling a passband signal at $\frac{1}{T}$ with $-f_c + k\frac{1}{T} \notin [f_c - W, f_c + W]$, we take the complex samples of the baseband signal.

$$\begin{aligned}
x_{BB} &= \left(e^{-i2\pi f_c t} x_{PB}(t) \right) \star \mathrm{LPF}_{W_c} && \tfrac{W}{2} \le W_c \le 2f_c - \tfrac{W}{2} \\
\Re(x_{BB}(t)) &= \left(x_{PB}(t) \cos(2\pi f_c t) \right) \star \mathrm{LPF}_{W_c} \\
Im(x_{BB}(t)) &= -\left(x_{PB}(t) \sin(2\pi f_c t) \right) \star \mathrm{LPF}_{W_c} \\
\Longrightarrow x_{BB}(t) &= \textstyle\sum_{-\infty}^{\infty} x_{BB}(l/W) \mathrm{sinc}(Wt - l) && \text{reconstruction of } x_{BB}
\end{aligned}$$

x_{BB} can be reconstructed from the samples taken once every $1/W$ seconds.

- *Reconstruction*: x_{PB} can be reconstructed from its baseband samples

$$x_{PB}(t) \quad = 2\Re\left(x_{bb}(t) e^{i2\pi f_c t} \right)$$

- *Energy*: $\|x_{BB}\|^2 = \frac{2}{W} \sum_{-\infty}^{\infty} |x_{BB}(l/W)|^2$
- *Inner Product*: $\langle x_{PB}, y_{PB} \rangle = \frac{2}{W} \Re \left(\sum_{-\infty}^{\infty} x_{BB}\left(\frac{l}{W}\right) y_{BB}^*\left(\frac{l}{W}\right) \right)$

8 Mapping Bits to Waveforms

Modulator The role of a modulator in a digital communication system is to *map the data bits into a signal* that is then fed into the channel.

8.1 From Bits to Real Numbers

$\{D_j\}$ incoming data bits
$\{X_l\}$ sequence of real numbers
Encoding mapping φ

$$\begin{aligned}\varphi : \{0,1\}^k &\rightarrow \mathbb{R}^n \\ \varphi : (d_1, \dots, d_k) &\mapsto (x_1, \dots, x_n)\end{aligned}$$

Rate $\frac{k}{n}$

Example "'two bits per real symbol"' Data bits are broken into pairs:
$\dots, (D_{-2}, D_{-1}), (D_0, D_1), (D_2, D_3), \dots$

$$(D_{2j}, D_{2j+1}) \mapsto \begin{cases} +3 & \text{if } D_{2j} = D_{2j+1} = 0 \\ +1 & \text{if } D_{2j} = 0 \text{ and } D_{2j+1} = 1 \\ -3 & \text{if } D_{2j} = D_{2j+1} = 1 \\ -1 & \text{if } D_{2j} = 1 \text{ and } D_{2j+1} = 0 \end{cases}$$

Example "'one-half bit per real symbol"'

$$D_j \mapsto \begin{cases} (+1, +1) & \text{if } D_j = 0 \\ (-1, -1) & \text{if } D_j = 1 \end{cases}$$

8.2 From Real Numbers to Waveforms through Linear Modulation

Linear Modulation Modulate k data bits $\{D_j\}_{j=1}^k$, map these to n real numbers $\{X_j\}_{j=1}^n$ and transmit the waveform

$$X(t) = A \sum_{l=1}^{n} X_l g_l(t)$$

- *Energy:* $\|X\|^2 = A^2 \sum_{l=1}^n \sum_{l'=1}^n X_l X_{l'} \langle g_l, g_{l'} \rangle$
  or for orthonormal $\{g_l\}$: $\|X\|^2 = A^2 \sum_{l=1}^n X_l^2$

Theorem Any linear modulator $X(t) = A \sum_{l=1}^n X_l g_l(t)$ is equivalent to a linear modulator $X(t) = A \sum_{l=1}^{n'} X_l' \phi_l(t)$ for some orthonormal functions $\{\phi_l\}$.

8.3 Recovering the Signal Coefficients with Matched Filters

- For some orthonormal $\{\phi_l\}$.

$$\begin{aligned} X_l &= \frac{1}{A} \left\langle X, \phi_l \right\rangle && l = 1, \dots, n \\ &= \frac{1}{A} (X \star \check{\phi}_l^*)(0) && \text{Matched Filter} \end{aligned}$$

- If they are time shifts of each other: $\phi_l = \phi(t - lT_s), \quad 1 \le l \le n$

$$X(t) = A \sum_{l=1}^{n} X_l \phi(t - lT_s)$$

$$X_l = \frac{1}{A} (X \star \check{\phi}_l^*)(lT_s)$$

9 The Nyquist Criterion

9.1 The Auto-correlation Function of a Deterministic Signal

Auto-correlation $\mathrm{R}vv(\tau)$ of a deterministic energy-limited signal v is

$$\mathrm{R}vv(\tau) = \|v\|^2 = \int_{-\infty}^{\infty} v(t+\tau) v^*(t) \, \mathrm{d}t, \qquad t \in \mathbb{R}$$

Properties of $\mathrm{R}vv$

Value at zero:	$\mathrm{R}vv(0) = \int_{-\infty}^{\infty} \lvert v(t) \rvert^2 \, \mathrm{d}t$
Maximum at zero:	$\lvert \mathrm{R}vv(\tau) \rvert \le \mathrm{R}vv(0)$
Conjugate symmetry:	$\mathrm{R}vv(-\tau) = \mathrm{R}^* vv(\tau)$
Integral representation:	$\mathrm{R}vv(\tau) = \int_{-\infty}^{\infty} \left\lvert \hat{v}(f) \right\rvert^2 e^{i2\pi f \tau} \, \mathrm{d}f$ $\hat{\mathrm{R}}vv(f) = \left\lvert \hat{v}(f) \right\rvert^2$
Uniform Continuity:	$\mathrm{R}vv(\cdot)$ is uniformly continuous
Convolution Representation:	$\mathrm{R}vv(\tau) = (v * \check{v}^*)(\tau)$

Self-similarity same as Auto-correlation
Shift-Orthogonality and Auto-Similarity

$$\int_{-\infty}^{\infty} \phi(t - lT_s) \phi^*(t - l'T_s) \, \mathrm{d}t = \mathrm{I}\{l = l'\} \qquad \Longleftrightarrow \qquad \mathrm{R}\phi\phi(lT_s) = \mathrm{I}\{l = 0\}$$

9.2 The Nyquist Criterion

Nyquist Pulse $v : \mathbb{R} \mapsto \mathbb{C}$ is a Nyquist Pulse of parameter $T_s > 0$ iff either

(a) Condition in time domain:

$$v(lT_s) = \mathrm{I}\{l = 0\}$$

(b) Nyquist Criterion: condition in frequency domain

$$\frac{1}{T_s} \sum_{j=-\infty}^{\infty} \hat{v}\left(f + \frac{j}{T_s}\right) = 1$$

Shift-Orthogonal Pulses
The complex functions $\phi(t - lT_s)$ are orthonormal $\left(\mathrm{R}\phi\phi(mT_s) = \mathrm{I}\{m = 0\}\right)$ iff

$$\frac{1}{T_s} \sum_{j=-\infty}^{\infty} \left\lvert \hat{\phi}\left(f + \frac{j}{T_s}\right) \right\rvert^2 = 1$$

Minimum Bandwidth of Shift-Orthogonal Pulses
If a function ϕ and its time shifts by non-zero integer multiples of T_s are orthonormal, then the bandwidth W of ϕ cannot be smaller than $1/(2T_s)$

$$W \ge \frac{1}{2T_s}$$

- *Equality* is achieved by any pulse ϕ for which

$$\left|\hat{\phi}(f)\right| = \sqrt{T_s} \cdot \mathrm{I}\left\{|f| < \frac{1}{2T_s}\right\}, \qquad f \in \mathbb{R} \qquad \text{"`Brickwall in f"'}$$

- In particular by the $\operatorname{sinc}(\cdot)$ pulse

$$\phi(t) = \frac{1}{\sqrt{T_s}} \operatorname{sinc}\left(\frac{t}{T_s}\right), \qquad t \in \mathbb{R}$$

Excess Bandwidth

Excess Bandwidth of a signal ϕ

$$100\% \cdot \left(\frac{\text{bandwidth of } \phi}{1/(2T_s)} - 1\right)$$

Band-edge symmetry

A real signal ϕ of excess bandwidth smaller than 100% and its time shifts are orthornomal iff

$$\left|\hat{\phi}\left(\frac{1}{2T_s} - f\right)\right|^2 + \left|\hat{\phi}\left(\frac{1}{2T_s} + f\right)\right|^2 = T_s, \qquad 0 \le f \le \frac{1}{2T_s}$$

Raised Cosine Spectra

$$\hat{\phi}(f) = \begin{cases} \sqrt{T_s} & \text{if } 0 \le |f| \le \frac{1-\beta}{2T_s} \\ \sqrt{\frac{T_s}{2}} \sqrt{1 + \cos\left(\frac{\pi T_s}{\beta}\left(|f| - \frac{1-\beta}{2T_s}\right)\right)} & \text{if } \frac{1-\beta}{2T_s} \le |f| \le \frac{1+\beta}{2T_s} \\ 0 & \text{if } |f| > \frac{1+\beta}{2T_s} \end{cases}$$

$$\phi(t) = \frac{2\beta}{\pi\sqrt{T_s}} \frac{\cos\left((1+\beta)\pi\frac{t}{T_s}\right) + \frac{\sin\left((1-\beta)\pi\frac{t}{T_s}\right)}{4\beta\frac{t}{T_s}}}{1 - \left(4\beta\frac{t}{T_s}\right)^2}, \qquad t \in \mathbb{R}$$

Roll-Off Factor β corresponds to the excess bandwidth $100\% \cdot \beta$

9.2.1 Nyquist Criterion for Passband Signals

The Nyquist Criterion still holds in the case of passband functions.

$W = 1/(2T_s)$ ***can be achieved if***

(a) $\left|\hat{\phi}(f)\right| = \sqrt{T_s}\mathrm{I}\left\{f_c - \frac{1}{4T_s} \le |f| \le f_c + \frac{1}{4T_s}\right\}$

(b) $4T_s f_c$ is odd

Baseband functions

- If we take a baseband Nyquist Pulse with bandwidth $W/2$ and baud $2T_s$ and multiply it by $\sqrt{2}\cos(2\pi f_c t)$ then we obtain a passband Nyquist Pulse with bandwidth W and baud T_s if $4T_s f_c$ is odd.

10 Energy and Power in PAM Signaling

10.1 PAM Signaling

IID independently and identically distributed

PAM signal generated from k iid random data bits $\{D_i\}$ taking the values $\{0, 1\}$ equiprobably, which are mapped by the encoder to n real symbols $\{X_i\}$

$$X(t) = A\sum_{l=1}^{n} X_l g(t - lT_s)$$

Baud T_s symbol-length

10.2 Energy in PAM Signaling

Expected Energy is

$$\begin{aligned} \mathrm{E} &= \mathrm{E}\left[\int_{-\infty}^{\infty} X^2(t)\,\mathrm{d}t\right] \\ &= A^2\mathrm{E}\left[\int_{-\infty}^{\infty}\left(\sum_{l=1}^{n} X_l g(t - lT_s)\right)^2 \mathrm{d}t\right] \end{aligned}$$

Energy per bit is $\mathcal{E}_b\left[\frac{\text{energy}}{\text{bit}}\right] = \frac{\mathrm{E}}{k}$

Energy per symbol is $\mathcal{E}_s\left[\frac{\text{energy}}{\text{symbol}}\right] = \frac{\mathrm{E}}{n}$

$$\begin{aligned} \mathcal{E}_s &= \mathrm{E}\left[\int_{\tau}^{\tau+T_s} X^2(t)\,\mathrm{d}t\right] = A^2\int_{\tau}^{\tau+T_s} \mathrm{E}\left[\left(\textstyle\sum_{-\infty}^{\infty} X_l g(t - lT_s)\right)^2\right]\mathrm{d}t \\ &= A^2\int_{\tau}^{\tau+T_s}\textstyle\sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty}\mathrm{E}\left[X_l X_{l'}\right] g(t - lT_s) g(t - l'T_s)\,\mathrm{d}t \end{aligned}$$

Power is $\mathcal{P}\left[\frac{\text{energy}}{\text{second}}\right] = \frac{\mathcal{E}_s}{T_s}$

Expected Energy with different $g(\cdot)$ ***and*** X_l

- $\phi(\cdot)$ orthonormal

$$\mathrm{E} = A^2\sum_{l=1}^{n}\mathrm{E}\left[X_l^2\right]$$

- X_l are zero mean and uncorrelated

$$\mathrm{E}\left[X_l\right] = 0, \qquad l = 1, \ldots, n \qquad \text{zero mean}$$

$$\mathrm{E}\left[X_l X_{l'}\right] = \begin{cases} \mathrm{E}\left[X_l^2\right] = \sigma_X^2 & \text{if } l = l' \\ 0 & \text{otherwise} \end{cases} \qquad \text{uncorrelated}$$

$$\mathrm{E} = A^2\sum_{l=1}^{n}\mathrm{E}\left[X_l^2\right]\|g\|^2 \qquad \text{energy}$$

- general case

$$\begin{aligned} \mathrm{E} &= A^2\sum_{l=1}^{n}\sum_{l'=1}^{n}\mathrm{E}\left[X_l X_{l'}\right] \mathrm{R}gg((l - l')T_s) \\ &= A^2\int_{-\infty}^{\infty}\sum_{l=1}^{n}\sum_{l'=1}^{n}\mathrm{E}\left[X_l X_{l'}\right] e^{i2\pi f(l-l')T_s}\left|\hat{g}(f)\right|^2\,\mathrm{d}f \end{aligned}$$

Special Expectations

$$\mathrm{E}\left[\left\{\pm\frac{d}{2},\pm\frac{3d}{2},\dots,\pm(2v-1)\frac{d}{2}\right\}\right]=\frac{d^2}{12}(2v+1)(2v-1)$$

10.3 The Power of a PAM Signal

Transmitted Power

$$\mathcal{P}=\lim_{T\to\infty}\frac{1}{2T}\mathrm{E}\left[\int_{-T}^{T}X^2(t)\,\mathrm{d}t\right]$$

Minimize the transmitted power

$$X'(t)=X(t)-\mathrm{E}\left[X(t)\right]$$

Power in PAM

- $\{X_l\}$ *Uncorrelated*: $\{X_l\}$ of zero mean, of variance σ_X^2, uncorrelated

$$\mathcal{P}=\frac{1}{T_s}A^2\sigma_X^2\left\|g\right\|^2$$

- $\{X_l\}$ *Stationary*: zero mean, with autocorrelation function $\mathrm{E}\left[X_lX_{l+m}\right]=\mathrm{K}_{XX}(m)$

$$\begin{aligned}\mathcal{P}&=\frac{1}{T_s}A^2\sum_{m=-\infty}^{\infty}\mathrm{K}_{XX}(m)\mathrm{R}_{gg}(mT_s)\\&=\frac{1}{T_s}A^2\int_{-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\mathrm{K}_{XX}(m)e^{i2\pi fmT_s}\left|\hat{g}(f)\right|^2\,\mathrm{d}f\end{aligned}$$

- *Orthonormality*: calculation is tricky, see script p 156.

$$\mathcal{P}=\frac{1}{T_s}A^2\lim_{L\to\infty}\frac{1}{2L+1}\sum_{L=-\infty}^{\infty}\mathrm{E}\left[X_l^2\right]$$

- *Block Mode*: see script p 158.

$$\begin{aligned}\mathcal{P}&=\frac{1}{nT_s}\int_{-\infty}^{\infty}\mathrm{E}\left[\left(A\sum_{l=1}^{n}X_lg(t-lT_s)\right)^2\right]\,\mathrm{d}t\\&=\frac{\mathrm{E}_n}{nT_s},\qquad \mathrm{E}_n=\int_{\tau}^{\tau+nT_s}\mathrm{E}\left[X^2(t)\right]\,\mathrm{d}t\end{aligned}$$

11 Power Spectral Density

Power Spectral Density $S_{XX}(\cdot)$ of a stochastic process X for any (nice) filter of impulse response h and frequency response $\hat{h}$ the average power at the filter's output when it is fed by X is given by

$$\text{Power in } X\star h=\int_{-\infty}^{\infty}S_{XX}(f)\left|\hat{h}(f)\right|^2\,\mathrm{d}f$$

For real stochastic processes X, $S_{XX}(\cdot)$ has to be symmetric.
Generate PSD out of $\sigma(f)$: $S_{XX}(f)=\frac{1}{2}\sigma(f)+\frac{1}{2}\sigma(-f)$

11.1 The PSD of PAM signals

(script p. 173)

$$\begin{aligned}Y(t)&=(X\star h)(t)\\&=A\int_{-\infty}^{\infty}X_l(g\star h)(t-lT_s)\end{aligned}$$

- $\{X_l\}$ *zero mean and uncorrelated*

$$S_{XX}(f)=\frac{A^2\sigma_X^2}{T_s}\left|\hat{g}(f)\right|^2$$

- $\{X_l\}$ *zero mean and stationary*

$$S_{XX}(f)=\frac{A^2}{T_s}\sum_{m=-\infty}^{\infty}\mathrm{K}_{XX}(m)e^{i2\pi fmT_s}\left|\hat{g}(f)\right|^2$$

- *Block Mode*:

$$S_{XX}(f)=\frac{A^2}{nT_s}\sum_{l=1}^{n}\sum_{l'=2}^{n}\mathrm{E}\left[X_lX_{l'}\right]e^{i2\pi f(l-l')T_s}\left|\hat{g}(f)\right|^2$$

12 Quadrature Amplitude Modulation

	PAM	*QAM*
rate	$R_s\left[\frac{\text{real symbols}}{\text{second}}\right]$	$\frac{R_s}{2}\left[\frac{\text{complex symbols}}{\text{second}}\right]$
baud	$T_s\geq\frac{1}{2W}$	$T_s\geq\frac{1}{W}$
bandwidth	$W\geq\frac{R_s}{2}$ [Hz]	$W\geq\frac{R_s}{2}$ [Hz] around f_c
communicate at	$2\left[\frac{\text{real dimensions/sec}}{\text{baseband Hz}}\right]$	$1\left[\frac{\text{complex dimension/sec}}{\text{passband Hz}}\right]$

12.1 A Complex PAM Baseband Representation

Baseband Representation

$$X_{BB}(t)=A\sum_{l=1}^{n}C_l\cdot g(t-lT_s)$$

Rate $\frac{k}{n}\left[\frac{\text{bits}}{\text{complex symbol}}\right]$
Real Passband Signal

$$\begin{aligned}X_{PB}(t)&=2\cdot\Re\left(X_{BB}(t)e^{i2\pi f_ct}\right)\\&=2A\sum_{l=1}^{n}\Re(C_l)\Re\left(g(t-lT_s)e^{i2\pi f_ct}\right)+2A\sum_{l=1}^{n}\Im(C_l)\Re\left(ig(t-lT_s)e^{i2\pi f_ct}\right)\\&=2A\sum_{l=1}^{n}\Re(C_l)g(t-lT_s)\cos(2\pi f_ct)-2A\sum_{l=1}^{n}\Im(C_l)g(t-lT_s)\sin(2\pi f_ct)\end{aligned}$$

Bandwidth If the puls shape g is bandlimited to $W/2$ Hz then the QAM signal will be bandlimited to W Hz around f_c.

Excess Bandwidth

$$100\% \cdot \left(\frac{\text{bandwidth of } \phi}{1/(2T_s)} - 1 \right)$$

12.2 Orthogonality Considerations

(script p. 183)

If the baseband puls shape ϕ satisfies (see Nyquist Criterion p. 10)

$$\int_{-\infty}^{\infty} \phi(t - lT_s)\phi^*(t - l'T_s)\, \mathrm{d}t = \mathrm{I}\{l = l'\}$$

$\Rightarrow T_s \geq \frac{1}{2 \cdot (W/2)}$ (script p. 186)

then the passband QAM signal

$$X_{PB} = \sqrt{2}A \sum_{l=1}^{n} \Re(C_l)\psi_{Q,l} + \sqrt{2}A \sum_{l=1}^{n} \Im(C_l)\psi_{I,l}$$

$$\psi_{Q,l}(t) = 2\Re\Big(\underbrace{\frac{1}{\sqrt{2}}\phi(t - lT_s)}_{\phi_{Q,l,BB}} e^{i2\pi f_c t} \Big)$$

$$\psi_{I,l}(t) = 2\Re\Big(\underbrace{i\frac{1}{\sqrt{2}}\phi(t - lT_s)}_{\phi_{I,l,BB}} e^{i2\pi f_c t} \Big)$$

12.3 Recovering $\{C_l\}$ via inner products

$$\Re(C_l) = \frac{1}{\sqrt{2}A} \left\langle X_{PB}, \psi_{Q,l} \right\rangle$$
$$\Im(C_l) = \frac{1}{\sqrt{2}A} \left\langle X_{PB}, \psi_{I,l} \right\rangle$$

Circuit $\left\langle r, \psi_{Q,l} \right\rangle$ and $\langle r, \psi_{I,l} \rangle$ for arbitrary $r(\cdot)$

$$s_{PB} = r \star \mathrm{BPF}_{W,f_c} \quad \text{Bandpassfilter}$$
$$s_{BB}(t) = s_{PB}(t) \cdot e^{-i2\pi f_c t} \star \mathrm{LPF}_{W_c}(t) \quad \text{Baseband Signal}$$
$$\Re(s_{BB}(t)) = s(t)cos(2\pi f_c t) \star \mathrm{LPF}_{W_c}(t)$$
$$\Im(s_{BB}(t)) = -s(t)sin(2\pi f_c t) \star \mathrm{LPF}_{W_c}(t)$$
$$\left\langle r, \psi_{Q,l} \right\rangle = \sqrt{2} \int \Re(s_{BB}(t)) \cdot \phi(t - lT_s)\, \mathrm{d}t \quad \text{Inner Product in Baseband}$$
$$\langle r, \psi_{I,l} \rangle = \sqrt{2} \int \Im(s_{BB}(t)) \cdot \phi(t - lT_s)\, \mathrm{d}t \quad \text{Inner Product in Baseband}$$

13 Energy, Power and PSD for QAM

13.1 Energy in QAM Signaling - Single Block Transmission

$$\mathrm{E} = \mathrm{E}\left[\int_{-\infty}^{\infty} X^2(t)\, \mathrm{d}t \right] = 2\mathrm{E}\left[\|X_{BB}\|^2 \right] = 2\mathrm{E}\left[\int_{-\infty}^{\infty} |X_{BB}(t)|^2\, \mathrm{d}t \right]$$

$$2\mathrm{E}\left[\|X_{BB}\|^2 \right] = 2\,|A|^2 \sum_{l=1}^{n} \sum_{l'=1}^{n} \mathrm{E}\left[C_l C_{l'}^* \right] \mathrm{R}gg\big((l' - l)T_s \big)$$
$$= 2\,|A|^2 \int_{-\infty}^{\infty} \sum_{l=1}^{n} \sum_{l'=1}^{n} \mathrm{E}\left[C_l C_{l'}^* \right] e^{i2\pi f (l'-l)T_s} \left| \hat{g}(f) \right|^2\, \mathrm{d}f$$

- *$\{C_l\}$ of zero mean and uncorrelated*

$$2\mathrm{E}\left[\|X_{BB}\|^2 \right] = 2\,|A|^2 \left\| g \right\|^2 \sum_{l=1}^{n} \mathrm{E}\left[|C_l|^2 \right]$$

- *time shifts of puls shape g orthonormal*

$$2\mathrm{E}\left[\|X_{BB}\|^2 \right] = 2\,|A|^2 \sum_{l=1}^{n} \mathrm{E}\left[|C_l|^2 \right]$$

13.2 The Power

- *In general*

$$\mathcal{P}_{PB} = \frac{\mathrm{E}_{PB}}{T_s} = \frac{2\mathrm{E}_{BB}}{T_s}$$

- *$\{C_l\}$ of zero mean an stationary*

$$\mathcal{P}_{PB} = \frac{2\,|A|^2}{T_s} \sum_{m=-\infty}^{\infty} \mathrm{K}_{CC}(m)\mathrm{R}gg(mT_s)$$
$$= \frac{2\,|A|^2}{T_s} \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \mathrm{K}_{CC}(m) e^{i2\pi f m T_s} \left| \hat{g}(f) \right|^2\, \mathrm{d}f$$

- *Power in Blockmode*

$$\mathcal{P}_{PB} = \frac{2|A|^2}{nT_s}\sum_{l=1}^{n}\sum_{l'=1}^{n}\mathrm{E}\left[C_l C_{l'}^*\right]\mathrm{R}gg\big((l'-l)T_s\big)$$
$$= \frac{2|A|^2}{nT_s}\int_{-\infty}^{\infty}\sum_{l=1}^{n}\sum_{l'=1}^{n}\mathrm{E}\left[C_l C_{l'}^*\right]e^{i2\pi f(l'-l)T_s}\left|\hat{g}(f)\right|^2\,\mathrm{d}f$$

13.3 The PSD

$$X_{BB}\star h'_{BB} = A\sum_{-\infty}^{\infty}C_l\cdot(g\star h'_{BB})(t-lT_s)$$

where $\hat{h}'_{BB}(f) = \hat{h}(f+f_c)\cdot \mathrm{I}\left\{|f|\le W/2\right\}$

- $\{C_l\}$ *of zero mean an stationary*

$$S_{XX}(f) = \frac{|A|^2}{T_s}\sum_{m=-\infty}^{\infty}\mathrm{K}_{CC}(m)e^{i2\pi(|f|-f_c)mT_s}\left|\hat{g}(|f|-f_c)\right|^2$$

- *Block mode*

$$S_{XX}(f) = \frac{|A|^2}{nT_s}\sum_{l=1}^{n}\sum_{l'=1}^{n}\mathrm{E}\left[C_l C_{l'}^*\right]e^{i2\pi(|f|-f_c)(l-l')T_s}\left|\hat{g}(|f|-f_c)\right|^2$$

14 The Univariate Gaussian Distribution

14.1 Standard Gaussian Random Variables

Standard Gaussian W has standard gassing distribution if its density is given by

$$f_W(w) = \frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}$$

Centered Gaussian $X = a\cdot W$

Gaussian $X = a\cdot W + b$ with $b = \mathrm{E}[X]$ and $a^2 = \mathrm{Var}(X)$

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

Transform to a standard Gaussian

$$X\sim\mathcal{N}(\mu,\sigma^2) \implies \frac{X-\mu}{\sigma}\sim\mathcal{N}(0,1)$$

Mathematical Notes

- Integral: $\int_{-\infty}^{\infty}e^{-w^2/2}\,\mathrm{d}w = \sqrt{2\pi}$
- Integral: $\sqrt{\frac{2}{\pi}}\int_{-\infty}^{\infty}e^{-2x^2}x^k\,\mathrm{d}x = \begin{cases}0 & k\text{ odd}\\ 1\cdot3\cdot5\cdots(k-1)\frac{1}{2^k} & k\text{ even}\end{cases}$
- Integral: $\int_{-\infty}^{\infty}e^{-\alpha x^2-\beta x}\,\mathrm{d}x = e^{\frac{\beta^2}{4\alpha}}\sqrt{\frac{\pi}{\alpha}}$

14.2 The Q-Function

Maps every $\alpha\in\mathbb{R}$ to the probability that a standard Gaussian random variable takes on a value that exceeds α

$$Q(\alpha) = \Pr[W>\alpha] = \frac{1}{\sqrt{2\pi}}\int_{\alpha}^{\infty}e^{-\xi^2/2}\,\mathrm{d}\xi \qquad \alpha\in\mathbb{R}$$

Probabilities

- $F_W(w) = \Pr[W\le w] = 1-Q(w)$
- $\Pr[a\le W\le b] = Q(a)-Q(b)$
- $\Pr[X>\alpha] = Q\left(\frac{\alpha-\mu}{\sigma}\right)$
- $Q(a)+Q(-a) = 1$
- $Q(0) = \frac{1}{2}$
- $Q(\alpha) = \frac{1}{\pi}\int_0^{\pi/2}e^{-\frac{\alpha^2}{2\sin^2\theta}}\,\mathrm{d}\theta \qquad \alpha\ge0$

Bounds

$$\frac{1}{\sqrt{2\pi}\alpha}e^{-\alpha^2/2}\left(1-\frac{1}{\alpha^2}\right) < Q(\alpha) < \frac{1}{\sqrt{2\pi}\alpha}e^{-\alpha^2/2} \qquad \alpha>0$$
$$Q(\alpha) < \frac{1}{2}e^{-\alpha^2/2} \qquad \alpha\ge0$$

14.3 Characteristics of Gaussians

Characteristic Function $\Phi_X(w)$ of X

$$\Phi_X(w) = \mathrm{E}\left[e^{iwX}\right] = \int_{-\infty}^{\infty}f_X(x)e^{iwx}\,\mathrm{d}x$$

Properties

- $\mathrm{E}[X^r] = \frac{\Phi_X^{(r)}(0)}{i^r}, \quad \text{if } \mathrm{E}[X^r]<\infty$
- $\Phi_X(\cdot)\equiv\Phi_Y(\cdot) \iff f_X(\cdot)\equiv f_Y(\cdot)$
- $\Phi_{X+Y}(w) = \Phi_X(w)\cdot\Phi_Y(w)$ if X and Y are independent random variables
- $X\sim\mathcal{N}(\mu,\sigma^2)\Rightarrow\Phi_X(w) = e^{iw\mu-\frac{1}{2}w^2\sigma^2}$

  $\mathrm{E}[W^v] = \begin{cases}1\cdot3\cdots(v-1) & \text{if } v \text{ is even}\\ 0 & \text{if } v\text{ is odd}\end{cases}, \qquad W\sim\mathcal{N}(0,1)$

  $\mathrm{E}[|W|^v] = \begin{cases}1\cdot3\cdots(v-1) & \text{if } v \text{ is even}\\ \sqrt{\frac{2}{\pi}}\cdot2^{(v-1)/2}\cdot\frac{v-1}{2}! & \text{if } v\text{ is odd}\end{cases}, \qquad W\sim\mathcal{N}(0,1)$

Moment Generating Function $M_X(\theta)$ of a random variable X

$$M_X(\theta) = \mathrm{E}\left[e^{\theta X}\right] = \int_{-\infty}^{\infty}f_X(x)e^{\theta x}\,\mathrm{d}x$$

Properties

- $X\sim\mathcal{N}(\mu,\sigma^2) \Rightarrow M_X(\theta) = e^{\theta\mu-\frac{1}{2}\theta^2\sigma^2}, \quad \theta\in\mathbb{R}$

The sum of independent Gaussians is a Gaussian
If $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ are independent

$$X + Y \sim \mathcal{N}(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$$

Linear combinations of independent Gaussians are Gaussian
If $X_1, \ldots, X_n$ are independent Gaussian random variables, then $Y = \sum_{l=1}^n \alpha_l X_l$ is Gaussian

$$\mathrm{E}[Y] = \sum_{l=1}^{n} \alpha_l \mathrm{E}[X_l], \qquad \mathrm{Var}(Y) = \sum_{l=1}^{n} \alpha_l^2 \mathrm{Var}(X_l)$$

$R^2 = X^2 + Y^2$ of two independent Gaussians is Rayleigh
If $X, Y \sim \mathcal{N}(0, \sigma^2)$

$$f_{R^2}(t) = \frac{1}{2\sigma^2} e^{-\frac{t}{2\sigma^2}}$$

The Projection of a Gaussian Random Vector onto a orthonormal Vector is Gaussian
If $Z \sim \mathcal{N}(\mathbf{0}, \sigma^2)$

$$\left\langle \mathbf{Z}, \phi \right\rangle = \tilde{Z} \sim \mathcal{N}(0, \sigma^2)$$

15 Binary Hypothesis Testing

15.1 Problem Formulation

Problem

- H random variable to guess
  - **Priors** $\pi_0 = \Pr[H = 0]$ and $\pi_1 = \Pr[H = 1]$
  - **non-degenerate** $\pi_o, \pi_1 > 0$
  - **degenerate** either of the priors is 0, guessing is deterministic
- Y observation

Optimal guessing rule

$$\phi^*_{\mathrm{Guess}}(y_{ops}) = \begin{cases} 0 & \text{if } \Pr[H = 0 \mid Y = y_{ops}] \geq \Pr[H = 1 \mid Y = y_{ops}] \\ 1 & \text{otherwise} \end{cases}$$

$$p^*(error \mid Y = y_{ops}) = \min\left\{\Pr[H = 0 \mid Y = y_{ops}], \Pr[H = 1 \mid Y = y_{ops}]\right\}$$

$$p^*(error) = \mathrm{E}\left[\min\left\{\Pr[H = 0 \mid Y], \Pr[H = 1 \mid Y]\right\}\right]$$

15.2 The Joint Law of H and Y

Y is a continuous random variable

$$f_Y(y) = \pi_0 f_{Y|H=0}(y) + \pi_1 f_{Y|H=1}(y) \qquad \text{unconditional density of } Y$$

$$\Pr[H = 0 \mid Y = y] = \frac{\pi_0 f_{Y|H=0}(y)}{f_Y(y)} \qquad \text{a-posteriori distribution}$$

$$\Pr[H = 1 \mid Y = y] = \frac{\pi_1 f_{Y|H=1}(y)}{f_Y(y)}$$

$$\phi^*_{\mathrm{Guess}}(y) = \begin{cases} 0 & \text{if } \pi_0 f_{y|H=0}(y) \geq \pi_1 f_{Y|H=1}(y) \\ 1 & \text{otherwise} \end{cases} \qquad \text{optimal guessing rule}$$

Y is a discrete random variable

$$P_Y(y) = \pi_0 P_{Y|H=0}(y) + \pi_1 P_{Y|H=1}(y) \qquad \text{Probability Mass Function}$$

$$\Pr[H = 0 \mid Y = y] = \frac{\pi_0 P_{Y|H=0}(y)}{P_Y(y)} \qquad \text{a-posteriori distribution}$$

$$\Pr[H = 1 \mid Y = y] = \frac{\pi_1 P_{Y|H=1}(y)}{P_Y(y)}$$

$$\phi^*_{\mathrm{Guess}}(y) = \begin{cases} 0 & \text{if } \pi_0 P_{y|H=0}(y) \geq \pi_1 P_{Y|H=1}(y) \\ 1 & \text{otherwise} \end{cases} \qquad \text{optimal guessing rule}$$

15.3 The MAP Decision Rule

The Maximum A-Posteriori decision rule

$$\phi_{\mathrm{MAP}}(y_{ops}) = \begin{cases} 0 & \text{if } \Pr[H = 0 \mid Y = y_{ops}] > \Pr[H = 1 \mid Y = y_{ops}] \\ 1 & \text{if } \Pr[H = 0 \mid Y = y_{ops}] < \Pr[H = 1 \mid Y = y_{ops}] \\ \text{equiprobably } \{0, 1\} & \text{if } \Pr[H = 0 \mid Y = y_{ops}] = \Pr[H = 1 \mid Y = y_{ops}] \end{cases}$$

$$\phi_{\mathrm{MAP}}(y_{ops}) = \begin{cases} 0 & \text{if } \pi_0 f_{Y|H=0}(y_{ops}) > \pi_1 f_{Y|H=1}(y_{obs}) \\ 1 & \text{if } \pi_0 f_{Y|H=0}(y_{ops}) < \pi_1 f_{Y|H=1}(y_{obs}) \\ \text{equiprobably } \{0, 1\} & \text{if } \pi_0 f_{Y|H=0}(y_{ops}) = \pi_1 f_{Y|H=1}(y_{obs}) \end{cases}$$

Likelihood Ratio Function $\mathbb{R}(\cdot)$

$$\mathrm{LR}(y) = \frac{f_{Y|H=0}(y)}{f_{Y|H=1}(y)}$$

with $\quad \alpha/0 = \infty, \quad \forall \alpha > 0 \quad$ and $\quad \frac{0}{0} = 1$

$$\phi_{\mathrm{MAP}}(y_{ops}) = \begin{cases} 0 & \text{if } \mathrm{LR}(y_{obs}) > \frac{\pi_0}{\pi_1} \\ 1 & \text{if } \mathrm{LR}(y_{obs}) < \frac{\pi_0}{\pi_1} \\ \text{equiprobably } \{0, 1\} & \text{if } \mathrm{LR}(y_{obs}) = \frac{\pi_0}{\pi_1} \end{cases}$$

Log Likelihood Ratio Function $LLR(\cdot)$

$$\text{LLR}(y) = \log \frac{f_{Y|H=0}(y)}{f_{Y|H=1}(y)}$$

with $\log \alpha/0 = \infty, \quad \forall \alpha > 0$ and $log \frac{0}{0} = 0$

$$\phi_{\text{MAP}}(y_{ops}) = \begin{cases} 0 & \text{if } \text{LLR}(y_{obs}) > \log \frac{\pi_0}{\pi_1} \\ 1 & \text{if } \text{LLR}(y_{obs}) < \log \frac{\pi_0}{\pi_1} \\ \text{equiprobably } \{0,1\} & \text{if } \text{LLR}(y_{obs}) = \log \frac{\pi_0}{\pi_1} \end{cases}$$

Probability of error

$$p^*(\text{error}) = \int_y \min\left\{\pi_0 f_{Y|H=0}(y_{ops}), \pi_1 f_{Y|H=1}(y_{obs})\right\} \, \mathrm{d}y$$

15.4 The ML Decision Rule

Is like the MAP rule, except that it ignores the priors.

$$\phi_{\text{ML}}(y_{ops}) = \begin{cases} 0 & \text{if } f_{Y|H=0}(y_{ops}) > f_{Y|H=1}(y_{obs}) \\ 1 & \text{if } f_{Y|H=0}(y_{ops}) < f_{Y|H=1}(y_{obs}) \\ \text{equiprobably } \{0,1\} & \text{if } f_{Y|H=0}(y_{ops}) = f_{Y|H=1}(y_{obs}) \end{cases}$$

$$\phi_{\text{ML}}(y_{ops}) = \begin{cases} 0 & \text{if } \text{LR}(y_{obs}) > 1 \\ 1 & \text{if } \text{LR}(y_{obs}) < 1 \\ \text{equiprobably } \{0,1\} & \text{if } \text{LR}(y_{obs}) = 1 \end{cases}$$

$$\phi_{\text{ML}}(y_{ops}) = \begin{cases} 0 & \text{if } \text{LLR}(y_{obs}) > 0 \\ 1 & \text{if } \text{LLR}(y_{obs}) < 0 \\ \text{equiprobably } \{0,1\} & \text{if } \text{LLR}(y_{obs}) = 0 \end{cases}$$

15.5 Processing of the Observations

Randomized Decision Rule can not outperform the optimal deterministic rule.

Conditionally Independent Random Variables
X and Y are conditionally independent given Z if

$$P_{X,Y|Z} = P_{X|Z}(x|z) \cdot P_{Y|Z}(y|z)$$

or, equivalently

$$P_{X|Y,Z}(x|y,z) = P_{X|Z}(x|z)$$
$$P_{Y|X,Z}(y|x,z) = P_{Y|Z}(y|z)$$

Processing W shall say that Z is the result of processing Y with respect to the random variable H and write $H \multimap Y \multimap Z$ if, when conditioned on Y, the random variables H and Z are independent. (Script p. 225)

15.6 The Bhattacharyya Bound

Some Bounds: $\min\{a,b\} \leq \sqrt{ab} \leq \frac{a+b}{2}$

$$p^*(\text{error}) \leq \frac{1}{2} \sum \sqrt{P_{Y|H=0}(y) P_{Y|H=1}(y)} = \frac{1}{2} \int \sqrt{f_{Y|H=0}(y) f_{Y|H=1}(y)} \, \mathrm{d}y$$

15.7 Sufficient Statistics

Sufficient Statistics Given: binary hypothesis testing problem of H based on the random vector $\mathbf{Y} \in \mathbb{R}^d$, $T : \mathbb{R}^d \to \mathbb{R}^{d'}$.
$\mathbf{T}(\mathbf{Y})$ forms a sufficient statistic for guessing H based on $\mathbf{Y}$ if, $\Pr[H = 0 \mid \mathbf{Y} = \mathbf{y}_{obs}]$ and $\Pr[H = 1 \mid \mathbf{Y} = \mathbf{y}_{obs}]$ can be computed from π_0, π_1 and $\mathbf{T}(\mathbf{Y})$.

Irrelevant Data Z is irrelevant data, if Y is a sufficient statistic for guessing H based on the pair (Y, Z). (Script p. 139)

The Factorization Theorem
$\mathbf{T}(\mathbf{Y})$ forms a sufficient statistic if the conditional densities can be factored in the form

$$f_{Y|H=0}(y) = g_0\big(\mathbf{T}(\mathbf{y})\big) \cdot h(y)$$
$$f_{Y|H=1}(y) = g_1\big(\mathbf{T}(\mathbf{y})\big) \cdot h(y)$$

15.8 Multi-dimensional Binary Hypothesis Testing

Problem

$$H = 0 : Y^{(j)} = s_0^{(j)} + Z^{(j)}, \quad j = 1, \ldots, \mathcal{J}, \qquad Z^{(j)} \sim \mathcal{N}(0, \sigma^2)$$
$$H = 1 : Y^{(j)} = s_1^{(j)} + Z^{(j)}, \quad j = 1, \ldots, \mathcal{J}, \qquad Z^{(j)} \sim \mathcal{N}(0, \sigma^2)$$

Ratios

$$\text{LR}(\mathbf{y}) = \frac{f_{\mathbf{Y}|H=0}(\mathbf{y})}{f_{\mathbf{Y}|H=1}(\mathbf{y})}$$
$$= \prod_{j=1}^{\mathcal{J}} \left(e^{-\frac{(y^{(j)} - s_0^{(j)})^2}{2\sigma^2} + \frac{(y^{(j)} - s_1^{(j)})^2}{2\sigma^2}} \right)$$
$$\text{LLR}(\mathbf{y}) = \frac{||\mathbf{s}_0 - \mathbf{s}_1||}{\sigma^2} \left(\left\langle \mathbf{y}, \phi \right\rangle - \frac{1}{2} \left(\left\langle \mathbf{s}_0, \phi \right\rangle + \left\langle \mathbf{s}_1, \phi \right\rangle \right) \right)$$
$$\phi = \frac{\mathbf{s}_0 - \mathbf{s}_1}{||\mathbf{s}_0 - \mathbf{s}_1||}$$
$$\tilde{\mathbf{Y}} = \left\langle \mathbf{Y}, \phi \right\rangle \qquad \text{is a sufficient statistic}$$
$$p^*(\text{error}) = \mathcal{Q}\left(\frac{||\mathbf{s}_0 - \mathbf{s}_1||}{2\sigma} \right)$$

16 Multi-Hypothesis Testing

$$\mathcal{M} = \{1, \dots, M\} \qquad \text{value set}$$
$$\pi_m = \Pr[M = m] \qquad \text{priors}$$

16.1 Optimal Guessing

1st approach

- Guess m_{guess} only iff

$$\Pr[M = m_{guess} \mid Y = y_{obs}] = \max_{m\in\mathcal{M}} \{\Pr[M = m \mid Y = y_{obs}]\}$$

- Error conditional on y_{obs}

$$p^*(\text{error} \mid Y = y_{obs}) = 1 - \left\{\max_{m\in\mathcal{M}} \Pr[M = m \mid Y = y_{obs}]\right\}$$

- Unconditional error

$$p^*(\text{error}) = 1 - \int_{y\in\mathcal{Y}} \left\{\max_{m\in\mathcal{M}} \Pr[M = m \mid Y = y_{obs}]\right\} f_Y(y)\, \mathrm{d}y$$

2nd approach by partitioning the space of outcomes into M sets $D_1, \dots, D_M$ where we guess m if $y \in D_m$. Guess m only if

$$\frac{\pi_m f_{Y|M=m}(y)}{f_Y(y)} = \max_{m'\in\mathcal{M}} \left\{\frac{\pi_{m'} f_{Y|M=m'}(y)}{f_Y(y)}\right\}$$

16.2 Example: 8PSK

$$\mathbf{Y} = \begin{pmatrix} Y^{(1)} \\ Y^{(2)} \end{pmatrix}, \qquad Y^{(1)} \sim \mathcal{N}(a_m, \sigma^2), \quad Y^{(2)} \sim \mathcal{N}(b_m, \sigma^2)$$
$$a_m = A\cos\frac{2\pi m}{8}, \quad b_m = A\sin\frac{2\pi m}{8}$$
$$f_{Y^{(1)},Y^{(2)}|M=m}(y^{(1)}, y^{(2)}) = \frac{1}{2\pi\sigma^2} e^{-\frac{(y^{(1)}-a_m)^2+(y^{(2)}-b_m)^2}{2\sigma^2}}, \qquad m \in \mathcal{M}$$

16.2.1 The "'Nearest Neighbor"' Decoding Rule

Guess m if

$$\left\|\mathbf{y} - \mathbf{s}_m\right\| = \min_{m'\in\mathcal{M}} \left\{\left\|\mathbf{y} - \mathbf{s}_{m'}\right\|\right\}$$

(See script p. 254)

16.2.2 Exact Error Analysis

Steps

1. Split $\mathbf{Y}$ up: $(Y^{(1)}, Y^{(2)}) = (-A, 0) + (Z^{(1)}, Z^{(2)})$ with $Z \sim \mathcal{N}(0, \sigma^2)$
2. Find the error region D_{error} conditional on y_{obs}
3. Find $f_{\mathbf{Z}}(\mathbf{z})$
4. $\int_{D_{error}} f_{\mathbf{Z}}(z)\, \mathrm{d}z$, change coordinates if needed

16.3 The Union Events Bound

Define for every $m' \neq m$ the set $\mathcal{B}_{m,m'}$

$$\begin{aligned}\mathcal{B}_{m,m'} &= \left\{y : \pi_{m'} f_{Y|M=m'}(y) \geq \pi_m f_{Y|M=m}(y)\right\} \\ &= \left\{y : \Pr[M = m'|Y = y] \geq \Pr[M = m|Y = y]\right\} \\ p^*(\text{error}) &\leq \sum_{m\neq m'} \int_{\mathcal{B}_{m,m'}} f_{Y|M=m}(y)\, \mathrm{d}y \\ &\leq \sum_{m\neq m'} \Pr[Y \in \mathcal{B}_{m,m'} \mid M = m] \qquad \text{union bound}\end{aligned}$$

Hint

$$p_{a,b}(\text{error}) = \mathcal{Q}\left(\frac{d_{a,b}}{2\sigma}\right) = \mathcal{Q}\left(\frac{\sqrt{\mathrm{E}_{a,b}}}{2\sigma}\right)$$

17 Gaussian Vectors

Gaussian Random Vector $\mathbf{X}$ if $\mathbf{X} - \mathrm{E}[\mathbf{X}]$ is a centered Gaussian.

Centered Gaussian has the same law as $A\mathbf{W}$ where A is a $n \times m$ Matrix and $\mathbf{W}$ is a vector whose components are IID $\mathcal{N}(0, 1)$

Notation

- $a^{(j,l)}$ or $[A]_{j,l}$ is the Row-j Column-l element
- $[A]_{j,l} = [A^T]_{l,j}$
- $(AB)^T = B^T A^T$

17.1 Expectations and Covariance Matrix

Expectations

$$\mathrm{E}[\mathbf{X}] = \begin{pmatrix} \mathrm{E}\left[X^{(1)}\right] \\ \vdots \\ \mathrm{E}\left[X^{(n)}\right] \end{pmatrix}$$

- $\mathrm{E}[AH] = A\,\mathrm{E}[H]$ where A is a deterministic and H a random matrix
- $\mathrm{E}[HB] = \mathrm{E}[H]\,B$ where B is a deterministic and H a random matrix
- $\mathrm{E}[H^T] = \mathrm{E}[H]^T$ where H is a random matrix

Covariance Matrix $\mathrm{K}_{\mathbf{XX}}$

$$\mathrm{K}_{\mathbf{XX}} = \mathrm{E}\left[\left(\mathbf{X} - \mathrm{E}[\mathbf{X}]\right)\left(\mathbf{X} - \mathrm{E}[\mathbf{X}]\right)^T\right]$$

$$= \mathrm{E}\left[\begin{pmatrix} X^{(1)} - \mathrm{E}\left[X^{(1)}\right] \\ \vdots \\ X^{(n)} - \mathrm{E}\left[X^{(n)}\right] \end{pmatrix}\begin{pmatrix} X^{(1)} - \mathrm{E}\left[X^{(1)}\right] & \cdots & X^{(n)} - \mathrm{E}\left[X^{(n)}\right] \end{pmatrix}\right]$$

$$= \begin{pmatrix} \mathrm{Var}(X^{(1)}) & \mathrm{Cov}[X^{(1)}, X^{(2)}] & \cdots & \mathrm{Cov}[X^{(1)}, X^{(n)}] \\ \mathrm{Cov}[X^{(1)}, X^{(1)}] & \mathrm{Var}(X^{(2)}) & \cdots & \mathrm{Cov}[X^{(2)}, X^{(n)}] \\ \vdots & & \ddots & \vdots \\ \mathrm{Cov}[X^{(n)}, X^{(1)}] & \mathrm{Cov}[X^{(n)}, X^{(2)}] & \cdots & \mathrm{Var}(X^{(n)}) \end{pmatrix}$$

- $\mathrm{K}_{\mathbf{YY}} = A\mathrm{K}_{\mathbf{XX}}A^T \qquad \mathbf{Y} = A\mathbf{X}$
- $\mathrm{K}_{\mathbf{XX}}$ is positive semi-definite
- $\mathrm{K}_{\mathbf{XX}}^T = \mathrm{K}_{\mathbf{XX}}$ symmetry
- $\mathrm{K}_{\mathbf{XX}}$ singular $\iff$ one of the components of X is a linear combination of the others
- Any positive semi-definite matrix is the covariance of some random vector.

17.1.1 Positive Semi-Definite Matrices

Positive Semi-Definite is a real $n \times n$ matrix K if

$$\alpha^T K \alpha \geq 0, \qquad \forall \alpha \in \mathbb{R}^n$$

iff one of the following equivalent conditions holds:

- $K = S^T S$ for some $n \times n$ matrix S ($S = \Lambda^{1/2} U^T$)
- All the eigenvalues of K are non-negative
- $K = U \Lambda U^T$ where
  Λ diagonal, containing the eigenvalues of K
  U U satisfies $UU^T = I_n$, U consists of orthonormal eigenvectors of K

Positive Definite is a real $n \times n$ matrix K if

$$\alpha^T K \alpha > 0, \qquad \forall \alpha \neq 0$$

iff one of the following equivalent conditions holds:

- $K = S^T S$ for some non-singular $n \times n$ matrix S ($S = \Lambda^{1/2} U^T$)
- All the eigenvalues of K are positive
- $K = U \Lambda U^T$ where
  Λ diagonal with positive diagonal entries
  U U is orthogonal: $UU^T = I_n$

Orthogonal $UU^T = U^T U = I_n$

Singular not invertible, $\det A = 0$

17.1.2 The Characteristic Function

If $\mathbf{X} \in \mathbb{R}^n$ then $\Phi_{\mathbf{X}}(\omega) : \mathbb{R}^n \to \mathbb{C}$

$$\Phi_{\mathbf{X}}(\omega) = \mathrm{E}\left[e^{i\omega^T \mathbf{X}}\right]$$
$$= e^{-\frac{1}{2}\omega^T \mathrm{K}_{\mathbf{XX}} \omega}$$
$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x}) e^{i \sum_{l=1}^{n} \omega^{(l)} x^{(l)}} \, \mathrm{d}x^{(1)} \cdots \mathrm{d}x^{(n)}$$

- X and Y have the *same distribution* if, and only if, they have identical characteristic functions
$$\left(\mathbf{X} \equiv \mathbf{Y}\right) \iff \left(\Phi_{\mathbf{X}}(\omega) = \Phi_{\mathbf{Y}}(\omega), \quad \forall \omega \in \mathbb{R}^n\right)$$
- X and Y are *independent* if, and only if
$$\mathrm{E}\left[e^{i(\omega_1 \mathbf{X} + \omega_2 \mathbf{Y})}\right] = \mathrm{E}\left[e^{i\omega_1 \mathbf{X}}\right] \cdot \mathrm{E}\left[e^{i\omega_2 \mathbf{Y}}\right] \qquad \omega_1, \omega_2 \in \mathbb{R}$$

17.2 A Standard Gaussian Vector

$$f_{\mathbf{W}}(\mathbf{w}) = \prod_{l=1}^{n} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{\left(w^{(l)}\right)^2}{2}}\right)$$
$$= (2\pi)^{-n/2} e^{-\frac{1}{2}\|\mathbf{w}\|^2}, \qquad \mathbf{w} \in \mathbb{R}^n$$
$$\mathrm{E}[\mathbf{W}] = 0, \qquad \mathrm{K}_{\mathbf{WW}} = I_n$$
$$\Phi_{\mathbf{W}}(\mathbf{w}) = e^{-\frac{1}{2}\|\mathbf{w}\|^2}, \qquad \mathbf{w} \in \mathbb{R}^n$$

17.3 Real Gaussian Random Vectors

$$\mathbf{X} = A\mathbf{W} + \mathbf{b}$$

- It the components of $\mathbf{X}$ are *independent scalar Gaussians* then $\mathbf{X}$ is a Gaussian vector.
- If $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent Gaussian Vectors, then $\begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}$ is Gaussian.
- If $\mathbf{X} = A\mathbf{W} + \mathbf{b}$ then
$$\mathrm{E}[\mathbf{X}] = \mathbf{b} \qquad \text{and} \qquad \mathrm{K}_{\mathbf{XX}} = AA^T$$
- If $\mathbf{X}$ is Gaussian
  - any subset of its components is Gaussian and
  - any permutation of its components is Gaussian.

17.4 The Mean and Covariance Determine the Law of a Gaussian

The mean and covariance or a Gaussian determine its law There is only one multivariate zero-mean Gaussian distribution of a given covariance matrix.

Accounting for the mean we have that there is *only one multivariate Gaussian distribution of a given mean vector and a given covariance matrix*.

The multivariate Gaussian distribution of mean μ and covariance K is denoted by $\mathcal{N}(\mu, K)$.

Uncorrelated $\text{Cov}[\mathbf{X}\,\mathbf{Y}] = 0$. In general: independent $\Longrightarrow$ uncorrelated

Jointly Gaussian $\mathbf{X}$ and $\mathbf{Y}$ are jointly Gaussian if $\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = A\mathbf{W} + \mathbf{b}$ is Gaussian. The density $f_{X,Y}(x, y)$ is of a Gaussian.
For jointly Gaussian vector holds: independent $\Longleftrightarrow$ uncorrelated

Properties of $K_{\mathbf{XX}}$

- If the components of the Gaussian random vector $\mathbf{X}$ are uncorrelated so that the *matrix* $K_{\mathbf{XX}}$ *is diagonal, then the components of* $\mathbf{X}$ *are independent*.
- If the components of a Gaussian random vector are independent in pairs, then they are independent.
- If $\mathbf{W}$ is a standard Gaussian and U is orthogonal, then $U\mathbf{W}$ is also a standard Gaussian vector.

17.5 A Canonical Representation of a Centered Gaussian

The centered Gaussian $\mathbf{X}$ with $\mathrm{K}_{\mathbf{XX}}$ has the same law as

$$U\Lambda^{1/2}\mathbf{W}$$

where

$\mathbf{W}$ is a standard Gaussian vector

U consists of the orthogonal eigenvectors of $\mathrm{K}_{\mathbf{XX}}$
U rotates the multivariate Gaussian $\Lambda^{1/2}\mathbf{W}$

Λ is diagonal and the diagonal entries correspond to the eigenvalues of $\mathrm{K}_{\mathbf{XX}}$

$$\Lambda^{1/2}\mathbf{W} = \begin{pmatrix} \sqrt{\lambda_1}W^{(1)} \\ \vdots \\ \sqrt{\lambda_n}W^{(n)} \end{pmatrix}$$ the v-th component is $\mathcal{N}(0, \lambda_v)$ distributed

Let $\mathbf{X} \sim \mathcal{N}(\mu, K)$ and U, Λ as above:

$$\Lambda^{-1/2}U^T(\mathbf{X} - \mu) \sim \mathcal{N}(\mathbf{0}, I)$$

17.6 The Density of a Gaussian Vector

$\mathbf{X} = U\Lambda^{1/2}\mathbf{W} = B\mathbf{W}$

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) &= \frac{f_{\mathbf{W}}(B^{-1}\mathbf{x})}{|\det(B)|} \\ &= \frac{\exp\left\{-\frac{1}{2}(\mathbf{x}-\mu)^T K^{-1}(\mathbf{x}-\mu)\right\}}{\sqrt{(2\pi)^n \det(K)}}, \qquad \mathbf{x} \in \mathbb{R}^n \end{aligned}$$

17.7 Linear Functions of Gaussian Vectors

$$\omega\mathbf{X} \sim \mathcal{N}(\omega^T\mu_{\mathbf{X}}, \omega^T\mathrm{K}_{\mathbf{XX}}\omega)$$

17.8 Review of the Multivariate Centered Gaussian Distribution

Multivariate Centered Gaussian See script p. 286.

18 Stochastic Processes

Stochastic Process real valued function of "'time"' t and "'luck"' ω:

$$\begin{aligned} X : \mathbb{R} \times \Omega &\to \mathbb{R} \\ (t, \omega) &\mapsto X(t, \omega) \end{aligned}$$

Sample Path stochastic process with a fixed experiment outcome ω

$$\begin{aligned} X(\cdot, \omega) : \mathbb{R} &\to \mathbb{R} \\ t &\mapsto X(t, \omega) \end{aligned}$$

Random Variable stochastic process with a fixed time instance t

$$\begin{aligned} X(t, \cdot) : \Omega &\to \mathbb{R} \\ \omega &\mapsto X(t, \omega) \end{aligned}$$

Joint Distribution Function specifies a stochastic process

$$F_{X(t_1),\dots,X(t_k)}(x_1, \dots, x_k) = \Pr[X(t_1) \le x_1, \dots, X(t_k) \le x_k]$$

Density $f_{X(t_1),\dots,X(t_k)}(\cdot)$ of a stochastic Process

$$F_{X(t_1),\dots,X(t_k)}(x_1, \dots, x_k) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} f_{X(t_1),\dots,X(t_k)}(\xi_1, \dots, \xi_k)\, \mathrm{d}\xi_1 \dots \mathrm{d}\xi_k$$

Mean $\mathbf{E}\left[X(t)\right], \overline{X}(t)$
Covariance $\mathbf{K}_{XX}(t, u)$

$$\mathrm{K}_{XX}(t, u) = \mathrm{E}\left[\left(X(t) - \overline{X}(t)\right)\left(X(u) - \overline{X}(u)\right)\right]$$

Properties of Stochastic Processes

Gaussian If for all choices of epochs $t_1, \dots, t_k$ the set of RV's $X(t_1), \dots, X(t_k)$ is a jointly Gaussian set of RV's.
A Gaussian process is completely specified by its mean and covariance function.

Stationary $X(\cdot)$ is stationary if for all sets of epochs $t_1, \dots, t_k$ and for all shifts τ
$F_{X(t_1),\dots,X(t_k)}(x_1, \dots, x_k) = F_{X(t_1+\tau),\dots,X(t_k+\tau)}(x_1, \dots, x_k)$
and if a density exists: $f_{X(t_1),\dots,X(t_k)}(x_1, \dots, x_k) = f_{X(t_1+\tau),\dots,X(t_k+\tau)}(x_1, \dots, x_k)$

Wide Sense Stationary if
$\mathbf{E}\left[X(t)\right]$ is constant for all t
$\mathbf{K}_{XX}(t, u)$ is a function of only $t - u$
stationary $\Longrightarrow$ wide sense stationary

18.1 Linear Functionals

$$Y = \int_{-\infty}^{\infty} g(t)X(t)\,\mathrm{d}t \qquad \text{random variable}$$
$$\mathrm{E}[Y] = \int_{-\infty}^{\infty} g(t)\mathrm{E}[X(t)]\,\mathrm{d}t \qquad \text{mean}$$
$$\mathrm{Var}(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(t)\mathrm{K}_{XX}(t,u)g(u)\,\mathrm{d}u\,\mathrm{d}t \qquad \text{variance}$$

If $g(\cdot)$ is continuous and non-zero only over a finite interval and if $X(\cdot)$ is stationary with a continuous covariance function.

If** $X(\cdot)$ **is Gaussian then $Y(\cdot)$ is also Gaussian.

If** $X(\cdot)$ **is WSS also see section 18.4 on page 20.

$$\mathrm{Var}(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(t)\mathrm{K}_{XX}(t-u)g(u)\,\mathrm{d}u\,\mathrm{d}t = \int_{-\infty}^{\infty} g(t)(\mathrm{K}_{XX}\star g)(t)\,\mathrm{d}t$$
$$\mathrm{Cov}[Y,Z] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(t)\mathrm{K}_{XX}(t-u)h(u)\,\mathrm{d}u\,\mathrm{d}t, \qquad Z = \int_{-\infty}^{\infty} h(t)X(t)\,\mathrm{d}t$$

18.2 Power Spectral Density or Spectrum

Let $X(\cdot)$ be WSS with covariance $\mathrm{K}_{XX}(\cdot)$

$$S_{XX}(f) = \int_{-\infty}^{\infty} \mathrm{K}_{XX}(\tau)e^{-i2\pi f\tau}\,\mathrm{d}\tau \qquad \Longleftrightarrow \qquad S_{XX}(f) = \widehat{\mathrm{K}_{XX}}(f)$$

- $\mathrm{K}_{XX}(\cdot)$ and $S_{XX}(f)$ are real and symmetric
- $S_{XX}(f) \geq 0, \quad f \in \mathbb{R}$
- If $\mathrm{K}_{XX}(\cdot)$ is the inverse Fourier transform of a symmetric, non-negative function, then it is the covariance of some WSS random process $X(\cdot)$

18.3 White Gaussian Noise

Let $Y_i = \int W(t)g_i(t)\,\mathrm{d}t$ where $W(\cdot)$ is WGN of intensity $N_0/2$.

$$\mathrm{Var}(Y) = \frac{N_0}{2}\int g_i^2(t)\,\mathrm{d}t$$
$$\mathrm{Cov}[Y,Z] = \frac{N_0}{2}\int g_i(t)g_j(t)\,\mathrm{d}t$$
$$\mathrm{E}\left[Y_iY_j\right] = \frac{N_0}{2}\delta_{ij}, \qquad \{\phi_i(\cdot)\} \text{ set of orthonormal functions}$$

18.4 WSS Processes in the Frequency Domain

$$\mathrm{Var}(Y) = \int_{-\infty}^{\infty} g(t)\big(\mathrm{K}_{XX}\star g\big)(t)\,\mathrm{d}t = \int_{-\infty}^{\infty} \left|\hat{g}(f)\right|^2 S_{XX}(f)\,\mathrm{d}f$$
$$\mathrm{Cov}[Y,Z] = \int_{-\infty}^{\infty} g(t)\big(\mathrm{K}_{XX}\star h\big)(t)\,\mathrm{d}t = \int_{-\infty}^{\infty} \hat{g}(f)S_{XX}(f)\hat{h}^*(f)\,\mathrm{d}f$$

18.5 Filtering Stochastic Processes

$$Y(\cdot,\omega) = \int_{-\infty}^{\infty} X(\tau,\omega)h(\cdot-\tau)\,\mathrm{d}\tau$$

$X(\cdot)$ ***zero mean***

$$Y(t) = \int_{-\infty}^{\infty} X(\tau)h(t-\tau)\,\mathrm{d}\tau$$
$$\mathrm{K}_{YY}(t,u) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau)\mathrm{K}_{XX}(t-\tau,u-\sigma)h(\sigma)\,\mathrm{d}\tau\,\mathrm{d}\sigma$$
$$\mathrm{K}_{YY}(v) = \left[\check{h}\star h\star \mathrm{K}_{XX}\right](v), \qquad \text{for } X(\cdot) \text{ WSS}$$
$$\mathrm{S}_{YY}(f) = \left|\hat{h}(f)\right|^2 S_{XX}(f), \qquad \text{for } X(\cdot) \text{ WSS}$$

19 Detection in Additive White Gaussian Noise

Detection of a discrete random variable $M \in \mathcal{M} = \{1,\ldots,M\}$ with priors $\pi_m = \Pr[M=m]$ and observable $Y(t) = s_m(t) + Z(t)$ where $\{s_m\}$ are deterministic finite-energy bandlimited to W Hz signals and $Z(\cdot)$ is WGN with intensity $\mathrm{S}_{ZZ}(f) = N_0/2$

$\mathcal{S} = \mathrm{span}(\{s_m\}) = \mathrm{span}(\{\phi_m\})$ where $\{\phi_m\}$ are orthonormal.

$\mathbf{T}$ ***forms a sufficient statistics*** The projection of $\mathbf{Y}$ onto the linear subspace $\mathcal{S}$ spanned by the signals, forms a sufficient statistics.

$$\mathbf{T} = \left(\left\langle \mathbf{Y},\phi_1\right\rangle,\ldots,\left\langle \mathbf{Y},\phi_d\right\rangle\right)^T \qquad \text{Gaussian vector}$$
$$\left\langle \mathbf{Y},\phi_l\right\rangle \sim \mathcal{N}\left(\left\langle s_m,\phi_l\right\rangle,\frac{N_0}{2}\right)$$
$$\tilde{\mathbf{T}} = \left(\left\langle \mathbf{Y},g_1\right\rangle,\ldots,\left\langle \mathbf{Y},g_{d'}\right\rangle\right)^T \qquad \text{where } \mathcal{S}\subseteq \mathrm{span}\{g_1,\ldots,g_{d'}\}$$
$$\sim \mathcal{N}(\mu_m,K)$$
$$\mu_m = \mathrm{E}\left[\tilde{\mathbf{T}}\mid M=m\right] = (\langle s_m,g_1\rangle,\ldots,\langle s_m,g_{d'}\rangle)^T$$
$$K = \frac{N_0}{2}\begin{pmatrix} \langle g_1,g_1\rangle & \langle g_1,g_2\rangle & \cdots & \langle g_1,g_{d'}\rangle \\ \langle g_2,g_1\rangle & \langle g_2,g_2\rangle & \cdots & \langle g_2,g_{d'}\rangle \\ \vdots & & \ddots & \vdots \\ \langle g_{d'},g_1\rangle & \langle g_{d'},g_2\rangle & \cdots & \langle g_{d'},g_{d'}\rangle \end{pmatrix}$$

You find the proof for the binary antipodal case in the script on page 307.

20 The Front-End Filter

Received signal is $Y(t) = s_m(t) + N(t)$ where $N(t)$ is a stationary Gaussian Process of spectrum $\mathrm{S}_{NN}(\cdot)$, and the signals $\{s_m\}_{m=1}^M$ are all bandlimited to W Hz.

Then $\mathbf{Y}\star\mathbf{h}$ with $\hat{h}(f) = 1, \quad \forall\left|f\right| \leq W$ *forms a sufficient statistics*. (Proof on page 315 in the script)

20.1 Whitening Filter

See in the script on page 317.

$$\left|\hat{g}(f)\right|^2 = \frac{1}{S_{NN}(f)}$$

21 Discrete Time Hypothesis Testing and Heuristics for Continuous Time

21.1 Discrete-Time Signals

The signal u has the value $u[j]$ at time j and duration J.

$$\|u\|^2 = \sum_{j=1}^{J} u^2[j] \qquad \text{Norm}$$

$$\langle u, v\rangle = \sum_{j=1}^{J} u[j]\, v[j] \qquad \text{Inner Product}$$

$$\left(u|_{\mathcal{S}} \in \mathcal{S}\right) \text{ and } \left(\langle u - u|_{\mathcal{S}}, s\rangle = 0, \forall s \in \mathcal{S}\right) \qquad \text{Projection of } u \text{ onto } \mathcal{S}$$

$$u|_{\mathcal{S}} = \langle u, s_1\rangle\, s_1 + \cdots + \langle u, s_m\rangle\, s_m$$

21.2 Discrete-Time Signals Observed in IID $\mathcal{N}(0, \sigma^2)$ Noise

$Y[j] = s_m[j] + Z[j]$ with priors $\Pr(M = m) = \pi_m$.

$$f_{\mathbf{Y}|M=m}(\mathbf{y}) = \frac{1}{(2\pi\sigma^2)^{J/2}} e^{-\frac{\|\mathbf{Y}-\mathbf{s}_m\|^2}{2\sigma^2}}$$

Optimal Decision Rule

$$\begin{aligned} m &= \operatorname{argmax}\left\{\pi_{m'} f_{\mathbf{Y}|M=m'}(\mathbf{y})\right\} \\ &= \operatorname{argmin}\left\{\frac{\|\mathbf{Y} - \mathbf{s}_{m'}\|^2}{2\sigma^2} + \log\frac{1}{\pi_{m'}}\right\} \end{aligned}$$

Sufficient Statistics

$\mathcal{S} = \operatorname{span}\{s_1, \ldots, s_M\}$ has dimension d. $\mathbf{Y}|_{\mathcal{S}}$ is the projection of $\mathbf{Y}$ onto $\mathcal{S}$.

(a) $\mathbf{Y}|_{\mathcal{S}}$ is a sufficient statistic

(b) $\left\langle\mathbf{Y}, \phi_1\right\rangle, \ldots, \left\langle\mathbf{Y}, \phi_d\right\rangle$ for orthonormal $\{\phi_l\}$ form a sufficient statistic

(c) the random variables $\langle\mathbf{Y}, \mathbf{g}_1\rangle, \ldots, \langle\mathbf{Y}, \mathbf{g}_L\rangle$ form a sufficient statistic for any L signals such that $\{s_1, \ldots, s_M\} \subset \operatorname{span}\{g_1, \ldots, g_L\}$

21.3 The Conditional Law of the Sufficient Statistic

The sufficient statistic $\left\langle\mathbf{Y}, \phi_1\right\rangle, \ldots, \left\langle\mathbf{Y}, \phi_d\right\rangle$ for orthonormal $\{\phi_l\}$ and the more general sufficient statistic $\langle\mathbf{Y}, \mathbf{g}_1\rangle, \ldots, \langle\mathbf{Y}, \mathbf{g}_L\rangle$ where $\{\mathbf{s}_1, \ldots, \mathbf{s}_M\} \subset \operatorname{span}\{\mathbf{g}_1, \ldots, \mathbf{g}_L\}$.

$$\left\langle\mathbf{Y}, \phi_l\right\rangle = \left\langle\mathbf{s}_m, \phi_l\right\rangle + \sigma W^{(l)}, \qquad l = 1, \ldots, d \qquad W^{(l)} \sim \mathcal{N}(0, 1)$$

$$\langle\mathbf{Y}, \mathbf{g}_l\rangle = \langle\mathbf{s}_m, \mathbf{g}_l\rangle + \tilde{Z}^{(l)}, \qquad l = 1, \ldots, d$$

$\tilde{Z}^{(l)}$ are jointly Gaussian with covariance matrix K_{ZZ}

$$K_{ZZ} = \sigma^2 \begin{pmatrix} \langle g_1, g_1\rangle & \langle g_1, g_2\rangle & \cdots & \langle g_1, g_L\rangle \\ \langle g_2, g_1\rangle & \langle g_2, g_2\rangle & \cdots & \langle g_2, g_L\rangle \\ \vdots & & \ddots & \vdots \\ \langle g_L, g_1\rangle & \langle g_L, g_2\rangle & \cdots & \langle g_L, g_L\rangle \end{pmatrix}$$

21.4 Probability of Error

$$\Pr(\text{error} \mid M = m) \leq \sum_{m' \neq m} Q\left(\sqrt{\frac{\|\mathbf{s}_m - s_{m'}\|^2}{4\sigma^2}}\right)$$

Index