Notes on Radar
Signal Processing

Douglas Cochran

*Lectures given at Arizona State University, January-May 2011.

Acknowledgments: (1) Portions of these notes were contributed by S. Howard and B. Moran. (2) Most of the notes were written at Echo Coffee, Scottsdale.
Some history of radar

Hertz 1880s - Observed reflection of radio waves
Tesla ~1900 - Described potential radar concept
1900-1930 - Initial technological proofs of concept (ships, aircraft)
WWII - Rapid development, including "pulsed radar" and expansion from HF/VHF to microwave

Applications today

Military
Air traffic control
Navigation
Automotive traffic/safety
Space protection & awareness

Weather
Sports
Mining
Ocean surveillance

Recent algorithmic advances
(Enabled by software-driven receiver and transmitter capabilities)

Space-Time Adaptive Processing (STAP)
Synthetic aperture (SAR/ISAR)
Time-reversal
Multi-path exploitation
Over-the-horizon (OTHR)
Simple Monostatic Active Radar (cartoon)

Monostatic: Transmitter & receiver are co-located
Active: System capabilities based on capturing reflections of a transmitted signal; i.e., under control of the radar

Some Elements Affecting Operation & Performance

- Aspect - Scattering of the incident signal is highly dependent on the orientation of the scatterers to the signal source
- Static vs. dynamic - Transmit/receive platforms, targets/clutter, and channel
- Multipath - And multistatic
- Cluttered vs. free-space environment
- Wavelength relative to scatterer material & geometric properties
- Aperture - Continuous, sampled, synthetic
- Waveform - Including pulsing pattern
- Antenna - Including array configuration
- Processing algorithms
Elementary Signal Models

Soon, we will appeal to more sophisticated physical understanding to motivate target and clutter models and related concepts such as the radar range equation & RCS. We begin with a simple but useful set of signal models to motivate mathematical concepts that will be important in the course.

Static scenario, Single point scatterer

\[ y(t) = s x(t-r) + n(t) \]

\[ \text{Scatter} \quad \text{Propagation} \quad \text{Noise} \]

"Strength" Delay

\[ x(t) \]

\[ r \quad \text{"Range"} \]

\[ \hat{r} = \frac{2r}{c} \Rightarrow r = \frac{\hat{r}}{2} \]

Estimation of the physical parameter \( r \) is via \( \hat{r} \).
* $s$ is related to the RCS and is typically complex (i.e., affects both magnitude & phase of $x$). Later it may be a RV, but for now it is deterministic.

* $n$ is noise.

**Static Scenario, Discrete point scatterers**

$$y(t) = \sum_k s_k \cdot (t - \tau_k) + n(t)$$

$$= \left[ s \ast x \right](t) + n(t)$$

with $$s(t) = \sum_k s_k \delta(t - \tau_k)$$

- Scatterer ranges are $\frac{CT_k}{2}$.
- The scene model is convolutional (LTI).

Closely spaced scatterers cause overlapping echoes.
Static Scenario, Continuum of Scatters

\[ y(t) = \int_{\mathbb{R}} s(\tau) \times (t-\tau) \, d\tau + n(t) \]

\[ = [x \ast s](t) + n(t) \]

\[ s(t) \]

\[ \rightarrow t \]

Notes:
1. \(|s|\) is typically very small and is aspect dependent in real problems.
2. \( s = s_r + s_c \)
3. There are many generalizations of this basic model.

Dynamic Model, Doppler

Suppose:
1. A monostatic system
2. Transmitted signal energy concentrated in a narrow frequency band around \( f_0 \)
3. The distance between the transmitter and a scatterer is decreasing at velocity \( v \).

Classical Lorentz transformation gives the frequency of the received signal as
\[ S_L = \left( \frac{1 + \frac{v}{c}}{1 - \frac{v}{c}} \right) S_L \]

Since \( \frac{v}{c} \) is small,

\[ \Delta S_L = (1 + \frac{v}{c}) \left( 1 + \frac{v}{c} + \frac{v^2}{c^2} + \ldots \right) \Omega_x \]

\[ \approx \left( 1 + \frac{2v}{c} + \text{h.o.t.} \right) \Omega_x \]

\[ \approx \left( 1 + \frac{2v}{c} \right) \Omega_x \]

The corresponding "Doppler shift" is \( \alpha = \Omega_L - \Omega_x \)

\[ \approx \frac{2v}{c} \]

Example:

\[ \text{Radar} \quad \xrightarrow{\text{1000 m/s}} \quad \text{Target} \]

\[ \nu = +1000 \text{ m/s} \quad (\approx \text{mach 3}) \]

\[ c = 3 \times 10^8 \text{ m/s} \quad \nu = 10^6 \text{ Hz} \]

\[ \Rightarrow \alpha = \frac{2\nu}{c} \approx 6.6 \times 10^{-6} \cdot 10^6 \text{ Hz} \approx 10 \text{ Hz} \]

Notes: 1) The notion of Doppler shift is applicable to narrowband signals; in a wideband signal, different frequency components are shifted by different amounts.
2. The model

\[ y(t) = e^{2\pi i \left( \frac{2\nu}{c} \right) t} x(t-t) + \text{noise} \]

relies on the "narrowband assumption" or "narrowband approximation of the Doppler effect."

3. An LTI model of the scene is inadequate!

In a narrowband monostatic system, \( t \) informs about range and \( \alpha \) about radar-scatterer relative velocity.

\[ \downarrow \]

Estimation of \( t \) and \( \alpha \) is a key problem in RSP.
Review of Deterministic Signal Theory

Fourier Transform

Def: $L^2(\mathbb{R})$ is the complex vector space of functions $x: \mathbb{R} \rightarrow \mathbb{C}$ with

$$\int_{\mathbb{R}} |x(t)|^2 \, dt < \infty$$

i.e., complex-valued finite energy signals.

- Inner (scalar, dot) product

$$\langle x, y \rangle = \int_{\mathbb{R}} x(t)y^*(t) \, dt$$

- Norm

$$\|x\| = \sqrt{\langle x, x \rangle} = \left( \int_{\mathbb{R}} |x(t)|^2 \, dt \right)^{1/2} < \infty$$

Note: $L^2(\mathbb{R}^n)$ is defined analogously for complex functions of $n$ variables; c.f.

$$\int_{\mathbb{R}} \ldots \int_{\mathbb{R}} |x(t_1, \ldots, t_n)|^2 \, dt_1 \ldots dt_n < \infty$$
Def: The Fourier transform of \( x \in L^2(\mathbb{R}) \) is the complex-valued function \( \hat{x} \) with

\[
\hat{x}(f) = \int_{\mathbb{R}} x(t) e^{-2\pi if t} dt \tag{*}
\]

Remarks: 1. \( x(t) = \int_{\mathbb{R}} \hat{x}(f) e^{2\pi if t} df \) (Inversion)

If \( \hat{x} \) is obtained from \( x \) by \((*)\), then this inversion formula may not reproduce \( x(t) \) at every \( t \), but

\[
\int_{\mathbb{R}} |x(t) - \int_{\mathbb{R}} \hat{x}(f) e^{2\pi if t} df|^2 dt = 0
\]

2. Parseval's theorem

\[
\langle \hat{x}, \hat{y} \rangle = \langle x, y \rangle
\]

\[||\hat{x}||^2 = ||x||^2 \] (i.e., the Fourier transform preserves energy)

3. Cauchy-Schwarz inequality

\[|\langle x, y \rangle|^2 \leq ||x||^2 ||y||^2\]

\[
\Theta(x, y) = \cos^{-1}\left(\frac{|\langle x, y \rangle|}{||x|| \cdot ||y||}\right)
\]
4. Convolution (filtering and windowing)

\[ \hat{x} \hat{y} = \hat{x} \hat{y} \]

\[ \hat{x} y = \hat{x} * \hat{y} \]

5. Time and frequency shift

\[ y(t) = e^{2\pi i \alpha t} x(t) \Rightarrow \hat{y}(f) = \hat{x}(f - \alpha) \]
\[ y(t) = x(t - r) \Rightarrow \hat{y}(f) = e^{-2\pi i f r} \hat{x}(f) \]

6. Autocorrelation

\[ R_x(r) = \int_{\mathbb{R}} x(t) x^*(t-r) dt \Rightarrow \hat{R}_x(f) = |\hat{x}(f)|^2 \]

Resolution

Consider the window function \( w(t) = \begin{cases} 1 & |t| < \frac{1}{2} \\ 0 & \text{else} \end{cases} \)

Direct calculation gives \( \hat{w}(f) = T \text{sinc}(Tf) = \frac{\sin \pi Tf}{\pi f} \)

* The notational convention here is \( \text{sinc}(u) = \frac{\sin \pi u}{\pi u} \)
If a complex sinusoid \( s(t) = e^{2\pi i \omega_0 t} \) is windowed by \( w(t) \) to form a pulse \( x(t) = s(t)w(t) \) of duration \( T \), its Fourier transform is

\[
\hat{x}(f) = \left[ \hat{w} * \hat{s} \right](f)
\]

Since \( \hat{s}(f) = \delta(f-\omega_0) \), \( \hat{x}(f) = \hat{w}(f-\omega_0) \)

\[
= T \text{sinc}(T(f-\omega_0))
\]

Heuristically, the spectrum of \( s \) is spread by the truncation giving a main lobe width (unit amplitude) of \( \frac{\omega_s}{2} \). If two sinusoids are separated by \( \Delta < \frac{\omega_s}{2} \) in frequency, truncation results in overlapping main lobes. If \( \Delta < \frac{1}{T} \), the spectrum of the truncated signal will not have local maxima at the frequencies of the two sinusoids.

Typically, a block of \( T \) seconds of data is needed to resolve frequency components spaced \( \frac{1}{T} \text{ Hz} \) apart.
Example

Transmitter/Receiver

Target

\[ r = \frac{2 \times 1500 \text{ m}}{c} = \frac{3 \times 10^3 \text{ m}}{3 \times 10^8 \text{ m/s}} = 10^{-5} \text{ s} \]

Propagation time

With a sinusoidal pulse of this duration at 10^6 Hz, the Doppler shift is only \approx 7\text{Hz}

while the resolution is 10^5 \text{Hz}!
Signals in Time-Frequency Phase Space

For $x \in L^2(\mathbb{R})$, define

$$\rho_x(t) = \frac{|x(t)|^2}{\|x\|^2}$$

Note that $\rho_x(t) > 0$ for all $t \in \mathbb{R}$ and $\int_{\mathbb{R}} \rho_x(t) \, dt = 1$. Interpreting $\rho_x$ as a probability density motivates the following definitions:

- **Time mean** $M_t(x) = \int_{\mathbb{R}} t \rho_x(t) \, dt$

- **Time variance** $\sigma_t^2(x) = \int_{\mathbb{R}} (t - M_t(x))^2 \rho_x(t) \, dt > 0$

Similarly, defining $\hat{\rho}_x(f) = \frac{|\hat{x}(f)|^2}{\|x\|^2}$ yields

- **Frequency mean** $M_f(x) = \int_{\mathbb{R}} f \hat{\rho}_x(f) \, df$

- **Frequency variance** $\sigma_f^2(x) = \int_{\mathbb{R}} (f - M_f(x))^2 \hat{\rho}_x(f) \, df$

The Heisenberg uncertainty principle, which is a consequence of the Cauchy-Schwarz inequality, shows

$$\sigma_t(x) \sigma_f(x) \geq \frac{1}{16\pi^2}$$
Displacement Operator

Def: Given $r, \alpha \in \mathbb{R}$, the displacement operator

$$D(\alpha, r): L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$$

is defined by

$$D(\alpha, r)x(t) = e^{2\pi i \alpha t}x(t-r).$$

Notes:
1. $D(\alpha, r)$ is linear and
   $$\|D(\alpha, r)x\| = \|x\|.$$
2. $D(0, r)x(t) = x(t-r)$ — Pure time shift
   $D(\alpha, 0)x(t) = e^{2\pi i \alpha t}x(t)$ — Pure frequency shift

3. In general,
   $$\mathcal{M}_t(D(\alpha, r)x) = \mathcal{M}_t(x) + r \quad \{\text{Displaces}\}$$
   $$\mathcal{M}_f(D(\alpha, r)x) = \mathcal{M}_f(x) + \nu \quad \{\text{"center"}\}$$

while

$$\sigma_t(D(\alpha, r)x) = \sigma_t(x) \quad \{\text{Preserves}\}$$

$$\sigma_f(D(\alpha, r)x) = \sigma_f(x) \quad \{\text{"spread"}\}$$

4. From 3, signals of arbitrary time and frequency means are easy to obtain. But the uncertainty principle constrains variances.
Ambiguity Function

**Def.** The (auto) ambiguity function of \( x \in L^2(\mathbb{R}) \)

\[
A_x(\alpha, \tau) = \langle x, D(\alpha, \tau) x \rangle = \int_{\mathbb{R}} x(t) x^*(t-\tau) e^{-2\pi i \alpha t} dt
\]

**Properties:**

1. \[A_{D(u,v)x}(\alpha, \tau) = \int_{\mathbb{R}} e^{2\pi i ut} x(t-v) e^{-2\pi i \alpha(t-\tau)} e^{2\pi i v \tau} dt = \int_{\mathbb{R}} x(t-v) x^*(t-\tau) e^{-2\pi i \alpha t} e^{2\pi i \tau v} e^{2\pi i u t} dt = e^{2\pi i (uv-v\tau)} A_x(\alpha, \tau)\]

Shifting \( x \) in time and frequency affects the phase of \( A_x \), but not its magnitude.

2. \[
\|A_x\|_{L^2(\mathbb{R})^2} = \iint_{\mathbb{R}^2} |A_x(\alpha, \tau)|^2 d\alpha d\tau = \iint \left[ \int_{\mathbb{R}} |x(t) x^*(t-\tau) e^{-2\pi i \alpha t} dt|^2 \right] \left[ \int_{\mathbb{R}} |x(r) x^*(r-\tau) e^{-2\pi i \alpha r} dr|^2 \right] d\alpha d\tau
\]

\[
= \iint \iint x(t) x^*(t-\tau) x^*(r) x(r-\tau) e^{2\pi i \alpha (r-t)} dt dr d\alpha d\tau
\]
\[
= \iiint x(t)x^*(t-t')x^*(r-r') \left[ \int e^{2\pi i \alpha (r-t')} d\alpha \right] dt \, dr \, dt'
\]
\[
= \iint x(t)x^*(t-t') \left[ \int \int x^*(r)x(r-t') S(r-t') \, dr \right] dt \, dr'
\]
\[
= \iint x(t)x^*(t-t')x^*(t-r)x(t-r') \, dt \, dr = \|x\|_{L^2(\mathbb{R})}^4
\]

This is "Moyal's identity", i.e.,

\[
\|Ax\|_{L^2(\mathbb{R}^2)}^2 = \|x\|_{L^2(\mathbb{R})}^4
\]

The energy in a signal's ambiguity function is the square of the signal's energy.

3. \[Ax(0,r) = \int x(t)x^*(t-r) \, dt = R_x(r)\]

i.e., the "zero-Doppler slice" of \(Ax\) is the autocorrelation of \(x\).

**Examples (notional):**

1. \[x(t) = S(t) \implies \hat{x}(f) = 1 \forall f\]

\[A_x(u,v) = \int_{\mathbb{R}} S(t)S(t-u) e^{-2\pi i u \alpha} \, dt = \begin{cases} 0 \quad t \neq 0 \\ S(t) \quad t = 0 \end{cases}\]

\[= S(v) \forall \alpha\]
Using property 1 above,

\[ s(t-t_0) = D(0, t_0) \times (t) \]

so, if \( y(t) = s(t-t_0) \),

\[ A_y(\alpha, t) = A_{D(0, t_0)} \times (\alpha, t) = e^{-2\pi i\alpha t_0} s(t) \]

\[ x(t) = 1 \forall t \implies \hat{x}(f) = s(f) \]

\[ A_x(\alpha, \tau) = \int_{\mathbb{R}} 1 \cdot 1 \cdot e^{-2\pi i\alpha t} dt = s(\alpha) \]

And if \( y(t) = e^{2\pi i\alpha t} = D(f_0, 0) \times (t) \), property 1 gives

\[ A_y(\alpha, \tau) = e^{2\pi i\alpha f_0} s(\alpha) \]
Some Detection Theory

Hypotheses

\[ H_0 = \{ \text{The signal of interest is absent} \} \]
\[ H_1 = \{ \text{The signal of interest is present} \} \]

\[ H_0 \cap H_1 = \emptyset \]
\[ \Pr(H_0 \cup H_1) = 1 = \Pr(H_0) + \Pr(H_1) \]

Decisions

\[ \rightarrow H_1 = \{ \text{Detector decides the signal is present} \} \]
\[ \rightarrow H_0 = \{ \text{Detector decides the signal is absent} \} \]

Data \( D \) - Keep abstract for the moment.

Bayes' Decision Rule

The detector chooses \( H_1 \) if \( \Pr(H_1|D) > \Pr(H_0|D) \); otherwise, it chooses \( H_0 \).

\[ \rightarrow H_1 = \left\{ \frac{\Pr(D|H_1) \Pr(H_1)}{\Pr(D)} > \frac{\Pr(D|H_0) \Pr(H_0)}{\Pr(D)} \right\} \]

\[ = \left\{ \frac{\Pr(D|H_1) \Pr(H_1)}{\Pr(D)} > \frac{\Pr(D|H_0) \Pr(H_0)}{\Pr(D)} \right\} \]

\[ = \left\{ \frac{\Pr(D|H_1)}{\Pr(D|H_0)} \frac{\Pr(H_1)}{\Pr(H_0)} > 1 \right\} \]

Likelihood ratio \( \frac{\Pr(D|H_1)}{\Pr(D|H_0)} \) \quad \text{Ratio of "priors"}
Typically, the data has the form

\[ D = \{ \text{the RV } x \text{ is observed to take on the value } x \} \]

In this case,

\[ \Pr(\{x \leq x < x + \Delta\} | H_1) = \int_x^{x + \Delta} f_X(x | H_1) \, dx \approx f_X(x | H_1) \, \Delta \]

and

\[ \Pr(\{x \leq x < x + \Delta\} | H_0) \approx f_X(x | H_0) \, \Delta \]

So the likelihood ratio is

\[ \Lambda(x) \approx \frac{\Delta f_X(x | H_1)}{\Delta f_X(x | H_0)} = \frac{f_X(x | H_1)}{f_X(x | H_0)} \]

In the limit as \( \Delta \to 0 \), this expression is exact provided \( f_X \) is smooth. So the Bayesian test is

\[ H_1 \leftarrow \left\{ \frac{f_X(x | H_1)}{f_X(x | H_0)} > \frac{\Pr(H_0)}{\Pr(H_1)} \right\} \]

Note that the role of the priors is to establish a "detection threshold" for \( \Lambda(x) \). If, for example, \( \Pr(H_0) \approx 0 \), even small likelihood ratios will lead to the detector deciding in favor of \( H_1 \).
Example: $H_0: x \sim N(0, \sigma^2)$, $H_1: x \sim N(s, \sigma^2)$, $\Pr(H_0) = \Pr(H_1)$

\[ \rightarrow H_1 = \left\{ \frac{e^{-(x-s)^2/2\sigma^2}}{\sqrt{2\pi}\sigma} > 1 \right\} \]

\[ = \left\{ e^{(2sx-s^2)/2\sigma^2} > 1 \right\} \]

\[ = \left\{ 2sx - s^2 > 0 \right\} = \left\{ x > \frac{s}{2} \right\} \]

\[ f_x(x|H_0) \quad f_x(x|H_1) \]

\[ \rightarrow H_0 \quad \rightarrow H_1 \]

Performance Metrics

\[ P_d = \Pr(\rightarrow H_1 | H_1) \quad "Probability \ of \ (correct) \ detection" \]

\[ P_f = \Pr(\rightarrow H_1 | H_0) \quad "Probability \ of \ false \ alarm" \]

\[ P_e = \Pr(\rightarrow H_1 | H_0) \Pr(H_0) + \Pr(\rightarrow H_0 | H_1) \Pr(H_1) \]

\[ = P_f \cdot \Pr(H_0) + (1-P_d) \Pr(H_1) \quad "Probability \ of \ error" \]

Minimizing $P_e$ is often a reasonable performance criterion, but it requires knowing $\Pr(H_0)$ and $\Pr(H_1)$. 
Minimizing Probability of Error

\[ P_E = \Pr(H_0) \int_{H_1} f_x(x | H_0) \, dx \]
\[ + \Pr(H_1) \int_{H_0} f_x(x | H_1) \, dx \]
\[ = \int_{H_1} \Pr(H_0) f_x(x | H_0) \, dx \]
\[ + \int_{H_0} \Pr(H_1) f_x(x | H_1) \, dx \]

Since both integrands are non-negative, each \( x \in X \) should be assigned to \( \rightarrow H_1 \) if \( \Pr(H_1) f_x(x | H_1) > \Pr(H_0) f_x(x | H_0) \) and to \( \rightarrow H_0 \) otherwise; i.e.,
\[ x \in H_1 \text{ if and only if} \]
\[ \frac{f_x(x | H_1)}{f_x(x | H_0)} > \frac{\Pr(H_0)}{\Pr(H_1)} \]

Hence the Bayesian decision rule minimizes \( P_E \), but it requires known priors.
Neyman-Pearson Detection

The Neyman-Pearson (NP) lemma establishes that the decision rule

\[ H_1 = \left\{ \frac{f_{x}(x|H_1)}{f_{x}(x|H_0)} > T \right\} \]

where \( T \) is a threshold is optimal in the sense that it maximizes \( P_d \) for given \( P_f \).

To see this, partition the data space \( X \) into \( R_0 \) and \( R_1 \) such that the same \( P_f \) provided by the NP rule is achieved by the new rule; i.e.,

\[ P_f = \int_{H_0} f_{x}(x|H_0) \, dx = \int_{R_1} f_{x}(x|H_0) \, dx \]

Then

\[ \int_{H_1 \cap R_0} f_{x}(x|H_0) \, dx + \int_{H_1 \cap R_1} f_{x}(x|H_0) \, dx = \int_{R_0 \cap H_1} f_{x}(x|H_0) \, dx + \int_{R_1 \cap H_0} f_{x}(x|H_0) \, dx \]

and

\[ \int_{H_1 \cap R_0} f_{x}(x|H_0) \, dx = \int_{R_1 \cap H_0} f_{x}(x|H_0) \, dx \]
This equation comes from assuming the two decision rules give the same $P_d$. Now consider the detection probabilities:

$$P_d = \int f_x(x|H_1) \, dx \quad P_d^{\text{new}} = \int f_x(x|H_1) \, dx$$

From the NP definition of $\rightarrow H_1$:

$$\begin{align*}
\sigma \leq & \quad \int \left[ f_x(x|H_1) - T f_x(x|H_0) \right] \, dx - \int \left[ f_x(x|H_1) - T f_x(x|H_0) \right] \, dx \\
\rightarrow H_1 \cap H_0 & \quad R_1 \cap R_0
\end{align*}$$

Using $\sigma$,

$$\int f_x(x|H_1) \, dx \geq \int f_x(x|H_1) \, dx$$

Adding $\int f_x(x|H_1) \, dx$ to both sides gives

$$P_d \geq P_d^{\text{new}}$$

Detecting a known signal in AWGN:

Consider the problem of detecting a known signal vector $S \in \mathbb{R}^n$ in additive white
Gaussian noise; i.e.,

\[ H_1: \quad X = S + N \quad \sim N[0, \sigma^2 I] \]
\[ H_0: \quad X = \sim N \]

The likelihood ratio is

\[
\Lambda(X) = \frac{c \exp \left\{ -\frac{1}{2} (X-S)^T (\sigma^2 I)^{-1} (X-S) \right\}}{c \exp \left\{ -\frac{1}{2} X^T (\sigma^2 I)^{-1} X \right\}}
\]

\[
= \exp \left\{ \frac{1}{2\sigma^2} (X^T S + S^T X - S^T S) \right\}
\]

So the NP detector is

\[ \rightarrow H_1 = \{ S^T X > T \} = \{ \langle X, S \rangle > T \} \]

Notes:

\(1\) The signal energy \( \| S \|^2 \) and noise variance \( \sigma^2 \) are absorbed into \( T \), so the detector has the same form regardless of these values.

\(2\) The test statistic \( \rho = \langle X, S \rangle = \sum_{k=1}^{n} s_k x_k \)
is the length of the projection of the data vector into the signal subspace of $\mathbb{R}^n$.

$\mathcal{H}_0: \rho_n = S^T X = S^T N \Rightarrow \rho_n \sim \mathcal{N}[0, \sigma^2 \| S \|^2]$

$\mathcal{H}_1: \rho_n = S^T (S + N) \Rightarrow \rho_n \sim \mathcal{N}[\| S \|^2, \sigma^2 \| S \|^2]$

\[
\mathcal{P}_a = \Pr(\rho_n > T | \mathcal{H}_1) = \int_T^\infty f_{\rho_n}(\rho) \, d\rho = 1 - Q\left(\frac{T - \| S \|^2}{\sigma \| S \|}\right)
\]

\[
\mathcal{P}_\alpha = \Pr(\rho_n > T | \mathcal{H}_0) = \int_T^\infty f_{\rho_n}(\rho | \mathcal{H}_0) \, d\rho = 1 - Q\left(\frac{T}{\sigma \| S \|}\right)
\]
Detecting a Known Signal in AWGN (cont'd)

In the problem

\[ H_0 : \mathbf{X} = \mathbf{N} \quad \mathbf{X} \sim \mathbb{R}^n \text{ known} \]
\[ H_1 : \mathbf{X} = \mathbf{S} + \mathbf{N} \quad \mathbf{N} \sim \mathcal{N}[0, \sigma^2 \mathbf{I}] \]

The average signal power is \( \frac{\| \mathbf{S} \|^2}{n} \) \[ \Rightarrow \quad \text{SNR} = \frac{\| \mathbf{S} \|^2}{\| \mathbf{N} \|^2 / n} \]

The average noise power is \( \sigma^2 \).

To see that, for fixed \( n \), detector performance depends only on SNR, consider two signals \( S_1 \) and \( S_2 \), and two noise vectors \( N_1 \sim \mathcal{N}[0, \sigma_1^2 \mathbf{I}] \) and \( N_2 \sim \mathcal{N}[0, \sigma_2^2 \mathbf{I}] \). The detection statistics are
\[ \rho_1 = \left< \mathbf{X}, \mathbf{S}_1 \right> \quad \text{and} \quad \rho_2 = \left< \mathbf{X}, \mathbf{S}_2 \right> \quad \text{with} \]
\[ H_0 : \rho_1 \sim \mathcal{N}[0, \sigma_1^2 \| \mathbf{S}_1 \|^2] \]
\[ H_1 : \rho_2 \sim \mathcal{N}[\| \mathbf{S}_2 \|^2, \sigma_2^2 \| \mathbf{S}_2 \|^2] \]

For \( P_f \) to be the same in both cases,
\[ Q \left( \frac{T_1}{\sigma_1 \| \mathbf{S}_1 \|} \right) = Q \left( \frac{T_2}{\sigma_2 \| \mathbf{S}_2 \|} \right) \]

\[ \Rightarrow \quad \frac{T_1}{\sigma_1 \| \mathbf{S}_1 \|} = \frac{T_2}{\sigma_2 \| \mathbf{S}_2 \|} \quad \checkmark \]
For $P_d$ to also agree,

$$Q\left(\frac{T_1 - \|S_1\|^2}{\sigma_1 \|S_1\|}\right) = Q\left(\frac{T_2 - \|S_2\|^2}{\sigma_2 \|S_2\|}\right)$$

Using $\otimes$, this implies

$$\frac{\|S_1\|}{\sigma_1} = \frac{\|S_2\|}{\sigma_2} \Rightarrow \text{SNR}_1 = \text{SNR}_2$$

**Matched Filter**

Given $S = \begin{bmatrix} S_0 \\ \vdots \\ S_{n-1} \end{bmatrix}$, define $h(k) = \begin{cases} S_{n-1-k} & k=0, \ldots, n-1 \\ 0 & \text{else} \end{cases}$

Consider filtering the data stream $x(k)$ with the LTI filter defined by this impulse response.

$$y(k) = [x * h](k) = \sum_{m=0}^{n-1} h(m) x(k-m)$$

For each $k$, $y(k) = \langle x^{(k)}_\infty, S \rangle$ where

$$x^{(k)}_\infty = \begin{bmatrix} x(k-n+1) \\ \vdots \\ x(k) \end{bmatrix}$$
The filter outputs are the detection statistics for testing for the presence of $S$ in $x(t)$, step-by-step (assuming AWGN).

**Complex Representation of Signals**

Consider a signal of the form

$$s(t) = M(t) \cos \left[ 2\pi \Omega t + \phi(t) \right]$$

where $M(t) \geq 0$ and $\phi(t)$ are real signals that vary much more slowly than $\cos(\Omega t)$.

- $s$ is called a "phase and amplitude modulated (PAM) sinusoid"

- $\Omega$ is the carrier frequency in Hertz

- $M(t)$ is the amplitude modulation (in radar, this is usually just a rectangular function)

- $\phi(t)$ is the phase modulation ($\phi'(t)$ is the frequency modulation).
Using \( \cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b) \),

\[
\begin{align*}
\mathbf{s}(t) &= M(t) \cos \left[ \varphi(t) \right] \cos [2\pi f_0 t] \\
\mathbf{x}(t) &= \text{In-phase component of } \mathbf{s}(t) \\
\mathbf{y}(t) &= M(t) \sin \left[ \varphi(t) \right] \sin [2\pi f_0 t] \\
\mathbf{y}(t) &= \text{Quadrature component of } \mathbf{s}(t)
\end{align*}
\]

Notes:

1. If both \( x(t) \) and \( y(t) \) are known,

\[
M(t) = \sqrt{x^2(t) + y^2(t)}
\]

\[
\varphi(t) = \arctan \left( \frac{y(t)}{x(t)} \right)
\]

But knowing only one of them is not sufficient to recover either \( M(t) \) or \( \varphi(t) \).

2. Writing \( z(t) = x(t) + iy(t) \),

\[
M(t) = \mid z(t) \mid \quad \varphi(t) = \angle z(t)
\]

**Homodyne Receiver**

\[
\begin{array}{c}
\hat{x}(f) \\
\Uparrow
\end{array}
\quad \quad
\begin{array}{c}
\text{Magnitude spectrum} \\
\text{of } x(t) \cos(2\pi ft)
\end{array}
\]

\[
\begin{array}{c}
\quad \\
-\pi \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \qu
With \( x \) a lowpass signal (i.e., at baseband), the Fourier transform of \( x(t) \cos(2\pi f_0 t) \) is
\[
\frac{1}{2} \hat{x}(f + f_0) + \frac{1}{2} \hat{x}(f - f_0),
\]
which will be concentrated around \( \pm f_0 \), as pictured. Similarly, the Fourier transform of \( y(t) \sin(2\pi f_0 t) \) is
\[
-\frac{1}{2} \hat{y}(f + f_0) + \frac{1}{2} \hat{y}(f - f_0),
\]
the magnitude of which is also concentrated around \( \pm f_0 \).

\[
\text{Magnitude spectrum of } y(t) \sin(\omega t) \text{ as } \omega \to f_0.
\]

So \( x(t) \cos(2\pi f_0 t) \cdot \cos(2\pi f_0 t) \) has Fourier transform
\[
\frac{1}{2} \hat{x}(f + 2f_0) + \frac{1}{2} \hat{x}(f - 2f_0) + \frac{1}{4} \hat{x}(f + f_0) + \frac{1}{4} \hat{x}(f - f_0)
\]
and a lowpass filter recovers \( \frac{1}{2} x(t) \). Similarly, a lowpass filter recovers \( \frac{1}{2} y(t) \) from \( y(t) \sin(\omega t) \cdot \sin(\omega t) \).

On the other hand, \( y(t) \sin(2\pi f_0 t) \cdot \cos(2\pi f_0 t) \) has Fourier transform
\[
-\frac{i}{4} \hat{y}(f + 2f_0) + \frac{i}{4} \hat{y}(f - 2f_0) + \frac{i}{4} \hat{y}(f + f_0) + \frac{i}{4} \hat{y}(f - f_0),
\]
i.e., the baseband component cancels out. Similarly for \( x(t) \cos(2\pi f_0 t) \cdot \sin(2\pi f_0 t) \). Hence

```
x(t)[1 + \cos(4\pi f_0 t)] - 2y(t) \sin(4\pi f_0 t)
```

\[
\text{This structure is known as a homodyne receiver.}
\]
Complex Gaussian Noise

The complex random vector $\mathbf{Z}_i = X + iY$ is complex normal (complex Gaussian) if the real random vector $[X, Y]$ is jointly normal.

The mean of $\mathbf{Z}_i$ is

$$\mu_{\mathbf{Z}_i} = E[\mathbf{Z}_i] = E[X] + iE[Y] = \mu_X + i\mu_Y$$

Its covariance matrix is

$$\Sigma_{\mathbf{Z}_i} = E[(\mathbf{Z}_i - \mu_{\mathbf{Z}_i})(\mathbf{Z}_i - \mu_{\mathbf{Z}_i})^*]$$

where * denotes complex conjugate transpose (i.e., hermitian transpose). $\Sigma_{\mathbf{Z}_i}$ is hermitian symmetric ($\Sigma_{\mathbf{Z}_i}^* = \Sigma_{\mathbf{Z}_i}$) and non-negative definite.

For physical reasons, complex normal vectors in signal processing are usually assumed to be circularly symmetric, i.e.,

$$E[(\mathbf{Z}_i - \mu_{\mathbf{Z}_i})^T(\mathbf{Z}_i - \mu_{\mathbf{Z}_i})] = 0$$

This is an assumption; it does not follow from the other parts of the definition.
If \( Z \) is complex normal and circularly symmetric with mean \( \mu \) and covariance \( \Sigma \), the notation

\[
Z \sim CN[\mu, \Sigma]
\]

is often used.

**Detection of a Known Narrowband Signal in AWGN**

Consider a narrowband signal on a carrier

\[
s(t) = \text{Re} \left[ S(t) e^{j\omega t} \right]
\]

and narrowband white Gaussian noise *

\[
n(t) = \text{Re} \left[ N(t) e^{j\omega t} \right]
\]

A detector is to address the problem

\[
H_0 : X(t) = N(t) \quad \Rightarrow \quad H_0 : X(t) = N(t)
\]

\[
H_1 : X(t) = S(t) + N(t) \quad H_1 : X(t) = S(t) + N(t)
\]

(On carrier) \quad \text{(At baseband)}

The discretized and truncated problem at baseband is:

\[ H_0 : \mathbf{X} = \mathbf{N} \]
\[ H_1 : \mathbf{X} = \mathbf{S} + \mathbf{N} \]
\[ \mathbf{S} \in \mathbb{C}^m \]
\[ \mathbf{N} \sim \mathcal{CN}[\mathbf{0}, \sigma^2 \mathbf{I}_{m \times m}] \]

So,
\[
\begin{align*}
  f_{\mathbf{X}}(\mathbf{X} | H_0) &= \frac{1}{(2\pi\sigma^2)^m} \exp \left\{ -\frac{1}{2\sigma^2} \mathbf{X}^* \mathbf{X} \right\} \\
  f_{\mathbf{X}}(\mathbf{X} | H_1) &= \frac{1}{(2\pi\sigma^2)^m} \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{X} - \mathbf{S})^* (\mathbf{X} - \mathbf{S}) \right\}
\end{align*}
\]

and hence the likelihood ratio is

\[
\Lambda(\mathbf{X}) = \exp \left\{ \frac{1}{2\sigma^2} \left[ (\mathbf{X}^* \mathbf{S} + \mathbf{S}^* \mathbf{X}) - \mathbf{S}^* \mathbf{S} \right] \right\}
\exp \left\{ -\frac{1}{2\sigma^2} \| \mathbf{S} \|^2 \right\}
\]

Assuming \( S \) is known, the NP detector is

\[ \rightarrow H_1 = \left\{ \text{Re} \left< \mathbf{S}, \mathbf{X} \right> > T \right\}. \]

*Phase in Radar Returns*

Consider a static scenario with colocated transmitter, one scatterer, and no noise.
A baseband signal $s(t)$ is modulated onto a carrier at frequency $\nu_c$, yielding

$$D(\nu_c, 0) s(t)$$

This signal is propagated through the channel providing the echo return

$$a D(0, \frac{2\nu_c}{c}) D(\nu_c, 0) s(t)$$

where $a$ is the channel gain ($|a| < 1$) and $r$ is the range to the scatterer. Finally, this signal is returned to baseband

$$x(t) = D(-\nu_c, 0) a D(0, \frac{2\nu_c}{c}) D(\nu_c, 0) s(t)$$

So,

$$x(t) = D(-\nu_c, 0) a e^{2\pi i \nu_c t} s(t)$$

$$= D(-\nu_c, 0) a e^{2\pi i \nu_c (t - 2r/c)} s(t - \frac{2r}{c})$$

$$= a e^{-2\pi i rt} e^{2\pi i \nu_c (t - 2r/c)} s(t - \frac{2r}{c})$$

$$= a \exp(-4\pi i \frac{r}{c}) s(t - \frac{2r}{c})$$

The phase shift $\exp(-4\pi i \frac{r}{c})$ depends on the carrier frequency $\nu_c$ and also the target range $r$, which is typically unknown. For this reason, the phase of a radar return is generally unknown.
Detection of a Signal with Unknown Phase in AWGN

Since the phase of the received signal is generally not known, it is valuable to consider optimal detection of a signal with unknown phase.

\[ H_0 : \mathbf{X} = \mathbf{N} \]
\[ H_1 : \mathbf{X} = \mathbf{S} e^{j\theta} + \mathbf{N} \]
\[ \mathbf{N} \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 \mathbf{I}\right) \]
\[ \theta \text{ uniform on } [0, 2\pi] \]
\[ \mathbf{S} \in \mathbb{C}^m \text{ known} \]

\[ f_{\mathbf{X}}(\mathbf{X} | H_0) = \frac{1}{(2\pi\sigma^2)^m} \exp\left\{ -\frac{1}{2\sigma^2} \mathbf{X}^* \mathbf{X} \right\} \]

\[ f_{\mathbf{X}, \theta}(\mathbf{X}, \theta | H_1) = \frac{1}{(2\pi\sigma^2)^m} \cdot \frac{1}{2\pi} \exp\left\{ -\frac{1}{2\sigma^2} (\mathbf{X} - \mathbf{S} e^{j\theta})^* (\mathbf{X} - \mathbf{S} e^{j\theta}) \right\} \]

So,

\[ A_{\theta}(\mathbf{X}) \triangleq \frac{f_{\mathbf{X}, \theta}(\mathbf{X}, \theta | H_1)}{f_{\mathbf{X}}(\mathbf{X} | H_0)} \]

\[ = \frac{1}{2\pi} \exp\left\{ -\frac{1}{2\sigma^2} \| \mathbf{S} \|^2 \right\} \exp\left\{ \text{Re} \frac{\mathbf{X}^* \mathbf{S}}{\sigma^2} e^{j\theta} \right\} \]

Denote \( \frac{\mathbf{X}^* \mathbf{S}}{\sigma^2} = \mathbf{L} e^{j\Theta} \) (i.e., in polar coordinates).

Then,
\[ \Lambda_T^X(X) = \frac{1}{2\pi} \exp \left\{ -\frac{1}{2\sigma^2} \| S \|^2 \right\} \exp \left\{ -\cos(\theta + \psi) \right\} \]

Integrating out the density of \( \psi \) leaves the marginal density \( f_X^X(X|H_0) \) in the numerator. Hence

\[ \Lambda(X) = \int_0^{2\pi} \Lambda_T^X(X) \, d\psi \]

\[ = \exp \left\{ -\frac{1}{2\sigma^2} \| S \|^2 \right\} \frac{1}{2\pi} \int_0^{2\pi} e^{-L \cos \psi} \, d\psi \]

Doesn't depend on \( \theta \).

\[ = \exp \left\{ -\frac{1}{2\sigma^2} \| S \|^2 \right\} I_0(L) \]

where \( I_0(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{x \cos \psi} \, d\psi \) is the modified Bessel function of order zero. \( I_0(L) \) is monotonic in \( L \). Since \( L e^{i\theta} = \frac{1}{\sigma^2} x^* s \),

\[ L = \frac{|x^* s|}{\sigma^2} \]

and the NP detector is

\[ \rightarrow H_1 = \left\{ |\langle s, x \rangle| > T \right\} \]
Summary of Detection in AWGN

1. $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{N} \sim \mathcal{N}[0, \sigma^2 \mathbf{I}] \Rightarrow P = \langle \mathbf{x}, \mathbf{s} \rangle$

2. $\mathbf{x} \in \mathbb{C}^m$, $\mathbf{N} \sim \mathcal{CN}[0, \sigma^2 \mathbf{I}] \Rightarrow P = \text{Re} \langle \mathbf{x}, \mathbf{s} \rangle$

3. $\mathbf{x} \in \mathbb{C}^m$ with unknown phase, $\mathbf{N} \sim \mathcal{CN}[0, \sigma^2 \mathbf{I}] \Rightarrow P = |\langle \mathbf{x}, \mathbf{s} \rangle|$

Sampling in Pulse-Doppler Radar

Pulse-Doppler radar systems involve several different samplings of physical signals.

1. Pulsing ("slow time")

\[
\begin{align*}
\text{So} & \quad \text{Si} & \quad \text{Si} & \quad \text{So} & \quad t_{0} & \quad t_{0} + T & \quad t_{0} + T + L & \quad t_{0} + (N-1)T & \quad t_{0} + (N-1)T + L \\
& \quad t_{0} & \quad t_{0} + T & \quad t_{0} + T + L & \quad \text{T}
\end{align*}
\]

$T = \text{Pulse Repitition Interval (PRI)}$

$\frac{1}{T} = \text{Pulse Repitition Frequency in Hz (PRF)}$

$L = \text{Pulse duration}$
Notes on $L$:

a) $L$ must be short enough to allow the full pulse to be transmitted before the return arrives from the closest scatterer of interest; i.e.,

$$L < \frac{2R_{\text{min}}}{c}$$

Example: $R_{\text{min}} = 1.5 \text{Km} \Rightarrow L < \frac{3 \times 10^3}{3 \times 10^8} \approx 10 \mu\text{sec}$

b) Typically $L \ll T$. The "duty cycle" $\frac{L}{T}$ is often $< 10^{-2}$.

c) The minimum pulse bandwidth is on the order of $\frac{L}{T}$ Hz.

Notes on $T$:

a) $T$ is generally chosen to avoid "range aliasing"

At transmitter:

```
\begin{array}{c}
t \\
t+L \\
t+T \\
\end{array}
```

"Range gate"

At the receiver:

Pulse 1's return falls between $t + \frac{2R_{\text{min}}}{c}$ and $t + \frac{2R_{\text{max}} + L}{c}$.

Pulse 2's between $t + T + \frac{2R_{\text{min}}}{c}$ and $t + T + \frac{2R_{\text{max}} + L}{c}$.
Thus for disjoint returns, 

\[ t + \frac{2r_{\text{max}}}{c} + L < t + T + \frac{2r_{\text{min}}}{c} + L \Rightarrow T > \frac{2(r_{\text{max}} - r_{\text{min}})}{c} + L \]

where \( L \) is typically small.

(b) \( \frac{1}{T} \) Hz is the sample rate used in Doppler processing (discussed later). So 

\[ T < \frac{1}{10^{\lambda_{\text{max}}} \cdot 10} \]

where \( 10^{\lambda_{\text{max}}} \) is the largest Doppler shift anticipated.

Example: Two LEO satellites approach each other at \( v \approx 1.5 \times 10^4 \) m/s. One uses an X-band radar with \( f_0 = 10^{10} \) Hz.

\[ 10^{\lambda_{\text{max}}} = \frac{2v}{c} \cdot f_0 \approx 10^4 \text{ Hz (10 KHz)} \]

\[ \Rightarrow T < 10^{-4} \text{ sec (100 μsec)} \]

(2) Sampling within a Pulse ("fast time")

Consider a LTI target model with frequency response \( \hat{h}(f) \). Denote the pulse's spectrum by \( \hat{S}_0(f) \). Then the return signal's spectrum is 

\[ \hat{X}(f) = \hat{h}(f) \hat{S}_0(f) \]
So the bandwidth of $x$ is no greater than that of $s_0$, which determines the minimum sampling rate via the Shannon sampling theorem. Doppler and nonlinearities affect this, but it is still used as an approximation.
Range Estimation, Stationary Case

with One Scatterer

Recall that the auto ambiguity function for a signal $s \in L^2(\mathbb{R})$ is defined by

$$A_s(\alpha, \tau) = \int_{\mathbb{R}} s(t) s^*(t-\tau) e^{-2\pi i \alpha t} dt = \langle s, D(\alpha, \tau)s \rangle$$

Analogously, the cross ambiguity of $x$ and $s$ is

$$A_{xs}(\alpha, \tau) = \int_{\mathbb{R}} x(t) s^*(t-\tau) e^{-2\pi i \alpha t} dt = \langle x, D(\alpha, \tau)s \rangle$$

Note that $A_{sx}(\alpha, \tau) = A_{xs}(-\alpha, -\tau) e^{2\pi i \alpha \tau}$.

Consider the problem of estimating the range $r$ from the baseband transmitted signal $s$ and the baseband received signal $x(t) = as(t - \frac{2\tau}{c}) + p(t)$

where $|a| < 1$ and $p$ is $\mathcal{Z}_N$. An analog matched filter would produce

$$A_{xs}(0, \tau) = \int_{\mathbb{R}} x(u) s^*(u-\tau) du$$
\[
\frac{1}{\text{Re}} \sum_{k} \frac{1}{(k-\frac{r}{c}-\tau)+\nu} + \sum_{k} \frac{1}{(k-\frac{r}{c}-\tau)+\nu} \]

where \( R_s \) is the deterministic autocorrelation function of \( s \) and \( \nu \) is complex normal with mean zero and variance independent of \( \tau \).

So,

\[
E \left[ |A_{x_s}(0, \tau)|^2 \right] = F \left[ (a R_s \left( \frac{2c}{c^2} - \tau \right) \nu \right] + \nu^* \left[ a R_s \left( \frac{2c}{c^2} - \tau \right) \nu \right]^* \]
\[
= |a|^2 \left| R_s \left( \frac{2c}{c^2} - \tau \right) \right|^2 + a R_s \left( \frac{2c}{c^2} - \tau \right) \nu^* F[\nu] + E[\nu^* \nu^*] \]
\[
= |a|^2 \left| R_s \left( \frac{2c}{c^2} - \tau \right) \right|^2 + \text{var}(\nu) \]

and hence

\[
\hat{\tau} = \frac{c}{2} \arg\max_{\tau} |A_{x_s}(0, \tau)|^2 \]

is an unbiased estimator for \( r \).
Application in Pulse Systems

\[
s(t) = \sum_{n=0}^{N-1} s_n(t-nT)
\]

supported in \([0,L]\), \(n=0,\ldots,N-1\) and \(L < T\).

Then

\[
A_{x,s}(0,t) = \int_R x(t)s^*(t-t)dt = \int_R x(u+t)s^*(u)du
\]

\[
= \int_0^L s^*_n(u)x(u+t)du + \int_T^{T+L} s^*_n(u)x(u+t)du
\]

\[
+ \cdots + \int_{(N-1)T+L}^{(N-1)T} s^*_n(u)x(u+t)du
\]

\[
= \sum_{n=0}^{N-1} \int_0^L s_n^*(u)x(u+nT+t)du
\]

\[
= \sum_{n=0}^{N-1} \int_0^L s_n^*(u)x(u+nT+t)du
\]

If \(s_0 = s_1 = \cdots = s_{N-1}\),

\[
A_{x,s}(0,t) = \sum_{n=0}^{N-1} \int_0^L s_n^*(t)x(t+nT+t)dt = \sum_{n=0}^{N-1} A_{x,s_n}(0,nt+t)
\]
Remarks

1. The complex cross-ambiguity function with $\alpha = 0$ can be computed “return-by-return,” provided the system parameters and target range remain constant or vary in a predictable way.

2. Range can be estimated from a single pulse return (one “slow time” cell) if $r$ is restricted to a suitable range gate.

3. The summation over pulses provides “coherent gain,” i.e.,

$$A_{x_{s_0}}(0, nT+r) = \int_{\mathbb{R}} x(t) s_{s_0}^*(t-nT-r) dt$$

$$= \int_{0}^{1} s_{s_0}^*(u) x(u+nT+r) du$$

$$= \int_{0}^{1} s_{s_0}^*(u) [a_{s_0}(u- \frac{2r}{c} + r) + n(u)] du$$

With $r = \frac{2r}{c}$, $A_{x_{s_0}}(0, nT+r) \sim \text{CN}[\|s\|^2, \sigma^2]$.

But if $r$ is chosen so the supports of $s_{s_0}(u)$ and $s_{s_0}(u- \frac{2r}{c} + r)$ are disjoint,

$$A_{x_{s_0}}(0, nT+r) \sim \text{CN}[0, \sigma^2]$$
So the SNR of the peak value is \( \frac{\|a\|^{2} \|s_{0}\|^{4}}{\sigma^{2}} \).

If all \( N \) pulses are summed,

\[ \tau = \frac{2r}{c} \Rightarrow A_{x}(0, \tau) \sim CN \left[ N \|a\|^{2} \|s_{0}\|^{2}, No^{2} \right] \]

No overlap \( \Rightarrow A_{x}(0, \tau) \sim CN \left[ 0, No^{2} \right] \)

so the peak SNR is \( \frac{N^{2} \|a\|^{2} \|s_{0}\|^{4}}{No^{2}} = \frac{N^{2} |a|^{2} \|s_{0}\|^{4}}{\sigma^{2}} \),

\( N \) times higher than if only a single pulse is used.

4. If the phase varies unpredictably from pulse to pulse, only "incoherent gain" is achieved by summing over pulse returns. This is less than coherent gain, but generally still better than using only a single pulse return.

5. If the returned pulse fills partially within the matched filter support, a "sidelobe value" depending on the autocorrelation \( R_{s} \), is obtained for the mean of \( A_{x}(0, \tau) \).
Discrete-Time Interpretation of Range Estimation; Stationary, One-scatterer Case

Suppose $s_0$ and $x$ are both sampled with sampling interval $\Delta$.

$$\tilde{s}_0(l) = s_0(l\Delta), \quad l = 0, \ldots, \frac{L}{\Delta}$$

$$\check{x}(l) = x(l\Delta), \quad l \in \mathbb{Z}$$

Assume, for convenience, $\frac{L}{\Delta}$ and $\frac{T}{\Delta}$ are both integers. A crude approximation to $A_{x\check{s}}(0,T)$ is

$$\tilde{A}_{x\check{s}}(0,m) = \sum_{n=0}^{N-1} \sum_{l=0}^{\frac{L}{\Delta}} \tilde{s}_0(l) \check{x}(l+m+nT)$$

Example: Baseband signal bandwidth = 1 MHz

$$\Delta = 10^{-6} \text{ sec (Nyquist)}$$

$$\Rightarrow \text{Range bin size} = c\Delta \approx 3 \times 10^8 \text{ m/sec} \cdot 10^{-6} \text{ sec} = 300 \text{ m}$$

This may be adequate accuracy for some applications, but is clearly not for others.
Interpolation in Range Estimation

Suppose $\Delta$ is less than or equal to the Nyquist interval for $s_o(t)$. Then

$$s_o(t) \approx \sum_{k=0}^{1/\Delta} s_o(k\Delta) \text{sinc} \left( \frac{t-k\Delta}{\Delta} \right)$$

and

$$x(t) \approx \sum_{k} x(k\Delta) \text{sinc} \left( \frac{t-k\Delta}{\Delta} \right)$$

In these expressions,

$$\text{sinc}(u) = \frac{\sin \pi u}{\pi u}$$

These are approximations for at least two reasons:

(a) $s_o$ is time limited to $[0, L]$ and thus can be only approximately bandlimited.

(b) The summation would require infinitely many terms to be exact, even for an ideally bandlimited signal.

Using these approximations,
\[ A_{sX}(\rho, \tau) = \sum_{n=0}^{N-1} \int_{0}^{L} s_{n}^{*}(t) \times (t + nT + \tau) \, dt \]

\[ \approx \sum_{n} \sum_{l} \int_{0}^{L} s_{n}^{*}(lA) \, \text{sinc} \left( \frac{t-lA}{A} \right) \sum_{k} \times (kA) \, \text{sinc} \left( \frac{t+nT+\tau-kA}{A} \right) \, dt \]

\[ = \sum_{n} \sum_{j} \sum_{k} s_{n}^{*}(lA) x(kA) \int_{0}^{L} \text{sinc} \left( \frac{t-lA}{A} \right) \, \text{sinc} \left( \frac{t+nT+\tau-kA}{A} \right) \, dt \]

The interpolating functions

\[ \Phi_{jk}(\tau) = \int_{0}^{L} \text{sinc} \left( \frac{t-lA}{A} \right) \, \text{sinc} \left( \frac{t+nT+\tau-kA}{A} \right) \, dt \]

depend on \( j, k, \tau \) (and \( nT \) and \( A \)), but not on the signal samples. Thus \( A_{sX}(\rho, \tau) \) can be computed (approximately) off the sampling grid to improve range estimation accuracy. This is computationally expensive, but the interpolants can be computed off line.

**Doppler Estimation, Single Scatterer**

In the presence of Doppler,

\[ x(t) = D(-\Delta, 0) a D(\alpha, \tau) D(\delta, 0) s(t) + n(t) \]

\[ = a e^{2\pi i \alpha t} e^{2\pi i \delta t} s(t-\tau) + n(t) \]
This expression assumes \( r \) is approximately constant throughout the duration of \( s(t) \). Issues of "range migration" will be considered later.

For a given \( r \), consider an \( N \)-vector of slow-time samples of \( x(t) \) starting at time \( t_0 \); i.e.,

\[
\underline{x} = \begin{bmatrix}
    x(t_0) \\
    x(t_0 + T) \\
    \vdots \\
    x(t_0 + (N-1)T)
\end{bmatrix}
\]

The components of \( \underline{x} \) are

\[
x(t_0) = a \, e^{2\pi i a t_0} e^{2\pi i s t_0} s(t_0 - r) + \xi(t_0)
\]

\[
x(t_0 + T) = a \, e^{2\pi i a(t_0 + r)} e^{2\pi i s(t_0 + r)} e^{2\pi i r T} s(t_0 + T - r) + \xi(t_0 + T)
\]

\[
x(t_0 + nT) = a \, e^{2\pi i a t_0} e^{2\pi i s t_0} s(t_0 + nT - r) e^{2\pi i n r T} + \xi(t_0 + nT)
\]

Assuming \( s(t) = \sum_{n=0}^{N-1} s_0(t - nT) \) and \( t_0 - r \in [0, 1] \),

\[
s(t_0 + nT - r) = s_0(t_0 - r)
\]

So, denoting \( x_n = x(t_0 + nT) \) and
\[ A = a e^{2\pi i \alpha t} e^{2\pi i \delta t} s_0(t_0 - r), \]

\[
\begin{bmatrix}
1 \\
e^{2\pi i \alpha} \\
\vdots \\
e^{2\pi i (N-1)\alpha}
\end{bmatrix}
\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix} + 
\begin{bmatrix} \nu_0 \\ \nu_1 \\ \vdots \\ \nu_{N-1} \end{bmatrix}.
\]

i.e., \( x \) is a discrete sinusoid at frequency \( \alpha T \) Hz in AWGN. Hence the DFT modulus is peaked at \( \alpha T \) Hz, providing an estimator for \( \alpha \) as

\[
\hat{\alpha} = \arg \max_{\alpha} |\hat{X}(\alpha)|^2
\]

Remarks:

1. Zero-padding of \( x \) is used to apply length \(-2^m\) FFTs and to do frequency interpolation.
2. Windowing is sometimes used to improve sidelobe characteristics.
3. In range estimation, basic processing is done in fast time, and integration across slow time is used to improve SNR. In Doppler estimation, basic processing is in slow time and integration over fast time (changing to) is used to improve SNR.
Key Parameters in Fast-time & Slow-time Processing

<table>
<thead>
<tr>
<th>Fast Time</th>
<th>Slow Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample interval</td>
<td>$T_s$ (PRI)</td>
</tr>
<tr>
<td>Sample rate</td>
<td>$\frac{1}{T}$ Hz (Doppler Bandwidth)</td>
</tr>
<tr>
<td>Duration</td>
<td>$N T_s$</td>
</tr>
<tr>
<td># samples</td>
<td>$N$</td>
</tr>
<tr>
<td>Frequency</td>
<td>$\frac{1}{NT}$ Hz</td>
</tr>
<tr>
<td>resolution</td>
<td></td>
</tr>
<tr>
<td>Integration</td>
<td>$N$ PRIs</td>
</tr>
<tr>
<td></td>
<td>$L/A$ samples</td>
</tr>
</tbody>
</table>

**Doppler Bandwidth**

Given multiple targets at the same range but different radial velocities relative to the radar, Doppler bandwidth is defined at that range by the maximal difference between relative radial velocities:

![Diagram of Doppler bandwidth](image-url)
A stationary scene can entail a non-zero Doppler bandwidth due to motion of the radar platform and geometry:

Moving Target Indicator (MTI)

Recall the slow-time sample vector for a single scatterer with Doppler $\nu_0$:

$$X = A \begin{bmatrix} e^{2\pi i \nu_0 T} \\ \vdots \\ e^{2\pi i \nu_0 (N-1) T} \end{bmatrix} + \begin{bmatrix} x_0 \\ \vdots \\ x_{N-1} \end{bmatrix}$$

If $X$ is passed through a high-pass filter with frequency response $\hat{H}(\nu)$ having magnitude of the form
then the output signal $y$ has spectral density $S_y(\omega) = \hat{H}(\omega) S_x(\omega)$ and:

1. The DC (zero-Doppler) component of $x$ is attenuated by the filter. For a stationary radar, this filtering removes or sharply reduces returns from stationary scatterers.

2. $\|y\|^2$ is a useful statistic for determining whether a moving target is present.

3. With a moving radar, the stationary clutter band can be similarly excised by a band-pass filter for MTI.

**Example (Slow Movers):** Target speed $\approx 6$ m/s

\[ f = 10 \text{ GHz} = 10^{10} \text{ Hz} \quad \left| v_{\max} \right| \approx \frac{12 \text{ m/s}}{3 \times 10^8 \text{ m/s}} \times 10^{10} \text{ Hz.} \]

\[ \Rightarrow \frac{1}{NT} \approx 10^3 \text{ Hz} \Rightarrow NT \approx 1 \text{ ms} = 400 \text{ Hz.} \]
Effect of Multiple Scatterers

With no noise and multiple scatterers, the return signal is

\[ x(t) = D(-\infty, 0) \sum_{k=1}^{K} a_k D(\alpha_k, r_k) D(\beta, 0) s(t) \]

\[ = \sum_k \underbrace{a_k e^{2\pi i \alpha_k t}}_{\varphi_k(t)} e^{2\pi i \beta t} s(t-r_k) \]

\[ \varphi_k(t) \]

So,

\[ A_{xs}(u, v) = \int_{\mathbb{R}} x(t) s^*(t-v) e^{-2\pi i ut} dt \]

\[ = \sum_k A_{\varphi_k s}(u, v) \]

Further,

\[ A_{\varphi_k s}(u, v) = \int_{\mathbb{R}} \varphi_k(t) s^*(t-v) e^{-2\pi i ut} dt \]

\[ = a_k e^{2\pi i \alpha_k u} \int_{\mathbb{R}} e^{-2\pi i \alpha_k t} s(t-r_k) s^*(t-v) e^{-2\pi i ut} dt \]

\[ = a_k e^{2\pi i \alpha_k u} A_s(u-\alpha_k, v+r_k) \]

Thus,

\[ A_{xs}(u, v) = \sum_{k=1}^{K} a_k e^{2\pi i \alpha_k u} A_s(u-\alpha_k, v+r_k) \]
Examples:

- Target 1
- Target 2

Radar

LFM "Up-Chirp"

Cross-Ambiguity $A_{xs}$

LFM "Down-Chirp"

Auto-Ambiguity $A_s$
The array antenna consists of a collection of $M$ omni-directional antennas at positions $\{x_0, \cdots, x_{M-1}\}$.

Narrow band signal transmitted from a point in the far field of the array antenna.

$$s(x, t) = A(x, t) \exp i(k \cdot x - \omega_c t)$$

where \(k\) is the wave vector

$$k = \left( \frac{2\pi}{\lambda} \right) \hat{k}, \quad \lambda = \frac{2\pi c}{\omega_c}$$

\(\hat{k}\) a unit vector in the direction of propagation and \(\lambda\) the wave-length.
The narrow band approximation assumes array elements are close enough that \( A(x_j, t) \approx A(x_0, t) \) so that

\[
s(x_j, t) = A(x_0, t) \exp \left( i \mathbf{k} \cdot (x_0 + (x_j - x_0)) - \omega_c t \right)
= \exp \left( i \mathbf{k} \cdot (x_j - x_0) \right) s(x_0, t)
\]

The vector \( \mathbf{a}(\mathbf{k}) \) is called the \textit{steering vector}

\[
\mathbf{a}(\mathbf{k}) = \begin{pmatrix}
1 \\
\exp \left( i \mathbf{k} \cdot (x_1 - x_0) \right) \\
\vdots \\
\exp \left( i \mathbf{k} \cdot (x_{M-1} - x_0) \right)
\end{pmatrix}
\]

The vector signal collected by the array is

\[
\mathbf{y}(t) = \mathbf{a}(\mathbf{k}) s(x_0, t)
\]
\[ x_m = x_0 + \frac{\lambda}{2} m(1, 0, 0) \]

\[ k = \frac{2\pi}{\lambda} (\cos \theta, \sin \theta, 0) \]

so

\[ k \cdot (x_m - x_0) = \pi m \cos \theta \]

and the steering vector has components

\[ a(k)_m = \exp (\pi i m \cos \theta) \]
Suppose that the array is equi-spaced and linear

\[ x_m = x_0 + md \hat{\ell} \]

Signals arriving from different directions (different \( k \)) will not be distinguishable if \( a(k) = a(k') \), i.e.,

\[ \frac{2\pi md}{\lambda} (\hat{k} - \hat{k'}) \cdot \hat{\ell} = 2\pi n, \quad n \in \mathbb{Z} \]

or

\[ \frac{d}{\lambda} (\cos \theta - \cos \theta') \in \mathbb{Z} \]

OK if the spacing is less than half a wavelength (spatial Nyquist sampling)

\[ d \leq \frac{\lambda}{2} \]
Detection Problem: Detect a signal from direction $\hat{k}$ in the presence of uncorrelated Gaussian white noise.

\[ H_1 : \ y_n = a(k)s_n + \nu_n \]
\[ H_0 : \ y_n = \nu_n \]

for $n = 0, \cdots, N - 1$. Take $E(\nu_n\nu_{n'}^\dagger) = \sigma^2 I_M \delta_{n,n'}$

Under $H_1$ the probability density for $\{y_n, n = 0, \cdots, N - 1\}$ is

\[
p(\{y\} | \hat{k}, \{s\}, H_1) = \frac{1}{(\pi\sigma^2)^N} \exp \left( -\frac{1}{\sigma^2} \sum_{n=0}^{N-1} \|y_n - a(k)s_n\|^2 \right)
\]
Detection in White Noise (2)

All we assume about is it power $\Sigma^2$, i.e., $E(|s_n|^2) = \Sigma^2$ for all $n$ and use the maximum entropy distribution corresponding to that constraint as a prior:

$$p(\{s\}) = \frac{1}{(\pi \Sigma^2)^N} \exp \left(-\frac{1}{\Sigma^2} \sum_{n=0}^{N-1} |s_n|^2 \right)$$

Then

$$p(\{y\}|\hat{k}, H_1) = \int \cdots \int p(\{y\}|\hat{k}, \{s\})p(\{s\}) \prod_{n=0}^{N-1} d\text{Re}(s_n)d\text{Im}(s_n)$$

$$= \frac{1}{(\pi \sigma^2 (M \Sigma^2 + \sigma^2))^N} \exp \left(-\frac{1}{\sigma^2} \sum_{n=0}^{N-1} \|y_n\|^2 \right) \times \exp \left(\frac{1}{\sigma^2 (M + \sigma^2 / \Sigma^2)} \sum_{n=0}^{N-1} |a(\hat{k})^\dagger y_n|^2 \right)$$
The log-Likelihood ratio is:

\[
\log \frac{p(\{\mathbf{y}\}|\mathbf{k}, H_1)}{p(\{\mathbf{y}\}|H_0)} = \frac{1}{\sigma^2 (M + \sigma^2 / \Sigma^2)} \sum_{n=0}^{N-1} |\mathbf{a}(\mathbf{k})^\dagger \mathbf{y}_n|^2 - N \log \left( M\Sigma^2 / \sigma^2 + 1 \right)
\]

The important part of this is the statistic \( \|z\|^2 \), where \( z \) is the vector with components

\[
z_n = \mathbf{a}(\mathbf{k})^\dagger \mathbf{y}_n
\]

Ignoring noise for the moment, if a signal \( s \) arrives from direction \( \mathbf{k}_0 \) then

\[
\|z\|^2 = |\mathbf{a}(\mathbf{k})^\dagger \mathbf{a}(\mathbf{k}_0)|^2 \sum_{n=0}^{N-1} |s_n|^2
\]
Each linear combiner of the $M$ channels

$$z_n = w^\dagger y_n$$

$$\|z\|^2 = |w^\dagger a(k_0)|^2 \sum_{n=0}^{N-1} |s_n|^2 + \text{noise}$$

defines a beam with beam pattern

$$P_w(k) = |w^\dagger a(k)|^2$$

The beam pattern associated with the steering vector $a(k')$

$$P_{a(k')}(k) = |a(k')^\dagger a(k)|^2$$
Beam Patterns – Uniform Linear Array

Four element uniform linear array with half wavelength spacing.

Four element uniform linear array with full wavelength spacing.
The assumptions used in deriving the above detector usually don’t apply.
Let’s go back to the model

\[ y_n = a(k)s_n + \nu_n \]

Suppose we know the covariance of the zero mean noise \( E(\nu_n\nu_n^\dagger) = R \), then we can de-correlate the noise by multiplying through by \( R^{-1/2} \).

\[ R^{-1/2}y_n = R^{-1/2}a(k)s_n + \nu'_n, \quad E(\nu'_n\nu'_n) = I \]

We can now use a variation \((M \rightarrow a(k)^\dagger R^{-1}a(k), \sigma \rightarrow 1)\) of our previous result...
Coping with Interference (2)

**log-Likelihood ratio:**

\[
\log p(\{y\} | k, H_1) / p(\{y\} | H_0)
\]

\[
= \frac{1}{(a(k)^\dagger R^{-1} a(k) + 1/\Sigma^2)} \sum_{n=0}^{N-1} |a(k)^\dagger R^{-1} y_n|^2
\]

\[- N \log \left( a(k)^\dagger R^{-1} a(k) \Sigma^2 + 1 \right)\]

\[
= \frac{\sum_{n=0}^{N-1} |w_{opt}^\dagger y_n|^2}{w_{opt}^\dagger R w_{opt} \left( 1 + \frac{w_{opt}^\dagger R w_{opt}}{\Sigma^2} \right)} - N \log \left( \frac{\Sigma^2}{w_{opt}^\dagger R w_{opt}} + 1 \right)
\]

where

\[
w_{opt} = \frac{R^{-1} a(k)}{a(k)^\dagger R^{-1} a(k)}
\]
Suppose we know $E(\nu_n \nu_n^\dagger) = R$. Take a weight vector $w$ and compute the power of the beam-formed signal

$$P_w = E\left(|w^\dagger y_n|^2\right)$$

$$= |w^\dagger a(k)|^2 \Sigma^2 + w^\dagger Rw$$

Find $w$ such that

$$\min_w w^\dagger Rw \text{ such that } w^\dagger a(k) = 1$$
Introduce a Lagrange multiplier $\mu$ and minimize

$$w^\dagger R w - \mu (w^\dagger a(k) - 1) \quad (1)$$

Take the derivative w.r.t $w$ equating it to zero

$$Rw - \mu a(k) = 0 \quad \Longrightarrow \quad w = \mu R^{-1} a(k)$$

The condition $w^\dagger a(k) = 1$ is satisfied if

$$w_{opt} = \frac{R^{-1} a(k)}{a(k)^\dagger R^{-1} a(k)}$$

This is a minimum since the second derivative of (1) is $R$ which is positive definite and the minimum value is

$$w_{opt}^\dagger R w_{opt} = \frac{1}{a(k)^\dagger R^{-1} a(k)}$$
Back to the Likelihood Ratio

\[
\frac{1}{N} \sum_{n=0}^{N-1} |w_{\text{opt}}^\dagger y_n|^2 \quad \text{Sample power out from MVDR beamformer}
\]

\[
w_{\text{opt}}^\dagger R w_{\text{opt}} \quad \text{Predicted interference power passed through MVDR beamformer}
\]

\[
\frac{\Sigma^2}{w_{\text{opt}}^\dagger R w_{\text{opt}}} \quad \text{SINR after MVDR beamformer}
\]

\[
L.R. = N \frac{\sum_{n=0}^{N-1} |w_{\text{opt}}^\dagger y_n|^2 / N}{w_{\text{opt}}^\dagger R w_{\text{opt}} \left(1 + \frac{w_{\text{opt}}^\dagger R w_{\text{opt}}}{\Sigma^2}\right)} - N \log \left(\frac{\Sigma^2}{w_{\text{opt}}^\dagger R w_{\text{opt}}} + 1\right)
\]
Adaptive Beamforming – MPDR

In general the interference $\nu$ comes from clutter or jamming so its covariance in unknown. Suppose that we know $E\left(y_n^* y_n\right) = Q$ – if we don’t know this it is at least a quantity we can estimate from data.

Take a weight vector $w$ and compute the power of the beam-formed signal

$$P_w = \frac{1}{N} \sum_{n=0}^{N-1} |w^* y_n|^2 = w^* Q w$$

The MPDR minimizes the power of the beamformer output under the condition of distortionless response for the direction of interest:

$$\min_w w^* Q w \text{ such that } w^* a(k) = 1$$
Fortunately, we have already done a similar problem, so we can just write down the answer...

The MPDR weight vector is

\[ w_{\text{mpdr}} = \frac{Q^{-1}a(k)}{a(k)^\dagger Q^{-1}a(k)} \]

This corresponds to a minimum since \( Q \) is positive definite and the minimum value is

\[ w_{\text{mpdr}}^\dagger Q w_{\text{mpdr}} = \frac{1}{a(k)^\dagger Q^{-1}a(k)} \]
In general the interference \( \nu \) comes from clutter or jamming so its covariance is unknown. We don’t know \( R_c \), we don’t \( \hat{Q} \), but we can compute the sample covariance of the data \( \hat{Q} \).

The MPDR weight vector is

\[
\hat{\omega}_{\text{mpdr}} = \frac{\hat{Q}^{-1} a(k)}{\text{tr}(\hat{Q}^{-1} a(k)a(k)^\dagger)}
\]

This corresponds to a minimum since \( \hat{Q} \) is positive definite (mostly) and the minimum value is

\[
\hat{\omega}_{\text{mpdr}}^\dagger = 1/(a(k)\hat{Q}^{-1}a(k))
\]
ML Estimate of Covariance

Some facts from linear algebra:

1. Any positive semi-definite matrix $R$ can be written as $R = U D U^+$ with $D = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_1^2 \\ 0 & \sigma_2^2 \\ 0 & \ddots \\ \end{pmatrix}$.

The columns of the unitary matrix $U$ are the orthonormal eigenvectors of $R$.

2. The extreme points of the Rayleigh quotient $q_v = \frac{v^+ R v}{v^+ v}$ are the eigenvectors of $R$ and its value at the extreme points are the eigenvalues.
ML Estimate of Covariance

Model: We are given $N$ samples $\{y_n: n=0, \ldots, N-1\}$ from an $M$-dimensional complex Gaussian distribution with mean zero and unknown covariance $R$.

The joint pdf for the data is

$$p(\{y_n\} | R) = \frac{1}{T^N |R|^N} e^{-\sum_{n=0}^{N-1} y_n^+ R^{-1} y_n} e^{-N \text{Tr}(R^{-1} W)}$$

where

$$W = \frac{1}{N} \sum_{n=0}^{N-1} y_n y_n^+$$
**ML Estimate of Covariance**

Need to maximise

\[ l = \log p(y_n \mid R) = -N \log |R| - N \text{Tr}(R^{-1}W) \] (1)

over all non-negative matrices \( R \).

First write \( R = UDU^T \) with \( D = \begin{pmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_{M-1}^2 \end{pmatrix} \)

the maximise (1) over the \( \sigma_j \). Now

\[ l = -2N \sum_{j=0}^{M-1} \log \sigma_j - N \sum_{j=0}^{M-1} V_{jj} \sigma_j^2 + \text{const.} \] (2)

where

\[ V = U^T W U \]

so

\[ \frac{\partial l}{\partial \sigma_k} = -\frac{2N}{\sigma_k} + 2N V_{kk} \sigma_k^{-3} = 0 \Rightarrow \sigma_k^2 = V_{kk} \]
ML Estimate of Covariance

Substituting back into (2)

\[ l = -N \sum_{j=0}^{M-1} \log V_{jj} + \text{const.} \]

Now we need to maximise over the unitary matrix \( U \). So write \( U \) as

\[ U = (e_0, e_1, \ldots, e_{M-1}) \quad e_i^* e_j = \delta_{ij} \]

then

\[ l = -N \sum_{j=0}^{M-1} \log e_j^* W e_j + \text{const.} \]

This maximise if the \( e_j \) are the normalised eigenvectors of \( W \).

Thus the ML estimator of \( R \) is

\[ \hat{R} = W \]

Note that the ML estimator of \( R^{-1} \) is \( W^{-1} \).
MAP Estimate of Covariance

For small sample sizes (which is often the case in beamforming and STAP) this estimate performs poorly. The problem lies in the estimation of the small eigenvalues.

However, in the array detection problems if nothing is present but the signal of interest we know that \( R \) is just receiver noise in each of the channel which is usually well calibrated.

Prior distribution: \( \tilde{R} \) is Wishart distributed. The pdf for \( R \) is

\[
p(R) = \frac{\Gamma_2^{q-M} \rho^{2q} e^{-\text{Tr}(\sigma^2 I R^{-1})}}{c(N,q)}
\]

where

\[
c(N,q) = \frac{\Gamma(N)}{\Gamma(q-M+1)} \prod_{j=1}^{q-M} T(q-M+j)
\]

Each channel has noise power \( \sigma^2 \).
**MAP Estimate of Covariance**

The posterior density for $R$ is

$$p(R \mid \{y_n\}) = \frac{p(R) p(\{y_n\} \mid R)}{\int_{R>0} p(R) p(\{y_n\} \mid R) dR}$$

or

$$p(R \mid \{y_n\}) = \frac{C}{|R|^{N-q+M+2}} e^{-\text{Tr}\left(R^{-1}(NW + \sigma^2 I)\right)}$$

The MAP estimator is the $R$ which maximises $p(R \mid \{y_n\})$.

Proceeding as for the ML estimate,

$$\hat{R}_{\text{MAP}} = \frac{1}{N-q+M+2} \left(NW + \sigma^2 I\right)$$

Diagonal loading

MAP estimate of $R^{-1}$ has no 2
The Need for STAP

MTI radar modes distinguish targets from clutter using Doppler. But what if the radar is moving rapidly.

Need joint adaptivity in angle and Doppler
Doppler Adaptivity

Recall Doppler processing:

Transmit a pulse train of identical pulses

\[ x_n = A e^{2\pi i f_d n} + v_n \]

\( f_d \) in cycles/PRI
Doppler Adaptable

For detection in white noise

CACFAR (Cell Averaging Constant False Alarm Rate) Detectors:

Look at the range cells adjacent to the bin of interest to gain information about the noise/clutter and so improve detection.
Doppler Adaptivity

Consider the $k^{th}$ range bin

$$x_n^k = Ae^{2\pi if_d n} + y_n^k$$

Estimate covariance $R$ of time samples from adjacent range bins.

Write this as

$$x^k = A\tau(f_d) + y^k$$

$$\tau(f_d)_n = e^{2\pi if_d n}$$

Assuming that we know $R$, the likelihood ratio detector is

$$L.R. = \frac{1}{(\tau^* R^{-1} \tau + \frac{1}{\Sigma^2})} \left| \tau R^{-1} x^k \right|^2 + \log(\tau^* R^{-1} \tau + \frac{1}{\Sigma^2}) - \log \Sigma^2$$

Replace FT by $x \rightarrow w(f)^* x$ where

$$w(f) = \frac{R^{-1} \tau(f)}{\tau(f)^* R^{-1} \tau(f)}$$

Note that the FT is $x \rightarrow \tau(f)^* x$
Doppler Adaptivity

CACFAR ML estimate of $R$ is

$$\hat{R} = \frac{1}{K} \sum_{l = k-K/2}^{k+K/2} x_l^T x_l$$

Choices:

1. Need to choose $K$ large enough to get a good estimate but small enough that the clutter statistics don’t change too much. Diagonal loading helps.

2. May need to leave guard cells if the target response spreads over more than one range cell.
STAP

Return

Matched filter each channel with a single pulse

Range hypothesis

Slow time
STAP

For each range bin there are $M \times N$ samples, where $N$ is the number of pulses in a CPI. We stack them as

$$x = (x_{0,0}, x_{0,N-1}, x_{1,0}, \ldots, x_{1,N-1}, \ldots, x_{M-1,0}, \ldots, x_{M-1,N-1})$$

so

$$x = A \otimes (k, f) + \gamma$$

where

$$a(k, f) = a(k) \otimes \tau(f)$$

i.e.,

$$a(k, f)_{n + Nm} = a(k)_m e^{2\pi i fn} \quad n = 0, \ldots, N-1, \quad m = 0, \ldots, M-1$$

Kronecker product
STAP

Assuming that we know $R$, the likelihood ratio detector is

$$
L.R. = \frac{1}{|a(k,f)^{\dagger} R^{-1} x|} \left( a(k,f)^{\dagger} R^{-1} a(k,f) + \frac{1}{\Sigma^2} \right) + \log \left( a(k,f)^{\dagger} R^{-1} a(k,f) + \frac{1}{\Sigma^2} \right) - \log \Sigma^2
$$

MVDR processor $x \rightarrow w(k,f)^{\dagger} x$ where

$$
w(k,f) = \frac{R^{-1} a(k,f)}{a(k,f)^{\dagger} R^{-1} a(k,f)}$$
STAP

We can use CACFAR in range to estimate the covariance $R$ except that now $R$ is $(M \times N)$-dimensional.

Three problems:

1. We need enough range cells to estimate $R$.

2. Suppose the receive array has 1000 elements and we have a 64 pulse CPI then $R$ is a $64000 \times 64000$ matrix and we need to estimate and invert an $R$ for each range cell.

3. Model mismatch - when good assumptions go bad
Some More on Covariance Estimation

Let's go back to the optimum beamformer

\[ w = \alpha \mathbf{R}^{-1} \mathbf{a} \]

In terms of the eigenvectors and eigenvalues of \( \mathbf{R} \)

\[ \mathbf{R}^{-1} = \sum_{j=0}^{M-1} \frac{1}{\lambda_j} \mathbf{e}_j \mathbf{e}_j^T \quad \lambda_0 > \lambda_1 \cdots > \lambda_{M-1} \]

then

\[ w = \tilde{\alpha} (\lambda_{\text{min}} \mathbf{R}^{-1}) \mathbf{a} \]

\[ = \tilde{\alpha} (\mathbf{a} - (\mathbf{I} - \lambda_{\text{min}} \mathbf{R}^{-1}) \mathbf{a}) \]

\[ = \tilde{\alpha} (\mathbf{a} - \sum_{j=0}^{M-1} \frac{\lambda_j - \lambda_{\text{min}}}{\lambda_j} (\mathbf{e}_j^T \mathbf{a}) \mathbf{e}_j) \]

Opt. white noise beamformer

Remove interference directions for steering vector.
Some More on Covariance Estimation

Not enough training data

→ Reduce the number of parameters we need to estimate.

Prior information:

\[ R = R_K + \sigma^2 I \]

A few interferers

where \( R_K \) is a rank K psd matrix.

You might want to try to find the ML estimate for \( R_K \) and \( \sigma^2 \)
Some More on Covariance Estimation

Another approach is to look for even more structured covariance matrices. Suppose we expect that $R_{ij} = R_{i-j}$.

$$
R = \begin{pmatrix}
R_0 & R_1 & R_2 & \cdots & R_{M-1} \\
R_1^* & R_0 & R_1 & \cdots & R_{M-2} \\
R_2^* & R_1^* & R_0 & \cdots & R_{M-3} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
R_{M-1}^* & \cdots & \cdots & R_1^* & R_0
\end{pmatrix}
$$

This matrix has only $M$ parameters.

And so it goes...