Elements of Hypothesis Testing - 2

Prof. C. Emre Koksal

ECE 7001: Stochastic Processes, Detection, and Estimation

THE OHIO STATE UNIVERSITY

Outline

Neyman-Pearson Hypothesis Testing

- Bayesian (and minimax) requires the choice of suitable costs C_{ij}. In many scenarios, there is no obvious set of cost assignments.

A more natural optimization criterion:
Define $P_D(\delta) = P(\vec{Y} \in \Gamma_1 | H_1)$, and $P_F(\delta) = P(\vec{Y} \in \Gamma_1 | H_0)$

Note that $P(\text{misdetection}) = 1 - P_D$ and $P(\text{deciding } 0 | H_0) = 1 - P_F$

Neyman-Pearson (NP) decision problem:

$$\max_{\delta(\cdot)} P_0(\delta) \rightarrow \text{power of test}$$

$$s.t. \quad P_F(\delta) \leq \alpha \rightarrow \text{size of test}$$

Lagrange Multiplier Approach

Lagrangian: $J(\delta,\lambda) = (1 - P_D(\delta)) + \lambda(P_F(\delta) - \alpha)$

$$\Rightarrow \qquad \max_{\lambda \geq 0} \min J(\delta, \alpha)$$

$$J(\delta,\lambda) = \underbrace{\int_{\Gamma_0(\delta)} p(\vec{y}|H_1)\, d\vec{y}}_{1-P_D} + \lambda \left(\underbrace{1 - \int_{\Gamma_0(\delta)} p(\vec{y}|H_0)\, d\vec{y}}_{P_F} - \alpha \right)$$

$$= \lambda(1-\alpha) + \int_{\Gamma_0(\delta)} \left[p(\vec{y}|H_1) - \lambda\, p(\vec{y}|H_0) \right]\, d\vec{y}$$

$$\Rightarrow \qquad p(\vec{y}|H_1) - \lambda\, p(\vec{y}|H_0) \underset{0}{\overset{1}{\gtrless}} 0$$

$$\Rightarrow \qquad \frac{p(\vec{y}|H_1)}{p(\vec{y}|H_0)} \underset{0}{\overset{1}{\gtrless}} \lambda \qquad \rightarrow \text{LRT!}$$

Choose λ for which $P_F = \alpha$. The constraint $(P_F = \alpha)$ is active, because with LRT, P_F and P_D are both non-increasing functions of λ

$\Rightarrow$ decrease λ until $P_F = \alpha$

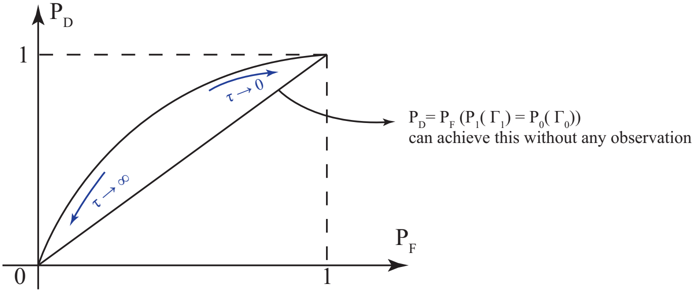
$\Rightarrow$ increase P_D as much as you can

$\Rightarrow$ choose λ such that

$$P(L(\vec{y}) > \lambda | H_0) = \alpha$$

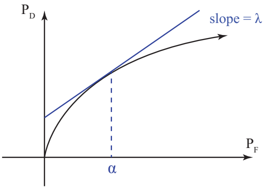
Receiver Operating Characteristics (ROC)

$P_D(\delta)$ vs. $P_F(\delta)$ plot. the associated decision rules δ are limited to the set of LRTs.$\Rightarrow$ $L(\vec{y}) \gtrless \tau$



$P_D - P_F$ curve (ROC)

1. Non-convex, non-decreasing (If convex at any point, we can achieve a better P_D at the same given P_F by some probabilistic decision rule)
2. Always above $P_D = P_F$ line (if not, reserve decision to obtaion a better performance)
3. if the pdf of $L(\vec{Y})$ exists for all $\vec{y}$, then $\left.\frac{\partial P_D}{\partial P_F}\right|_{P_F=\alpha} = \lambda$



Example:

$H_0 : Y \sim \mathcal{N}(\mu_0, \sigma^2)$, $H_1 : Y \sim \mathcal{N}(\mu_1, \sigma^2)$, $\mu_1 > \mu_0$

Recall LRT:

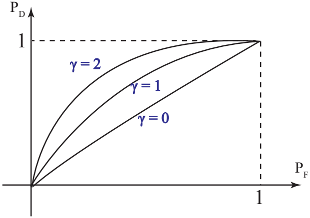
$$Y \underset{0}{\overset{1}{\gtrless}} Y_\lambda = \frac{\sigma^2}{\mu_1 - \mu_0} \ln \lambda + \frac{\mu_1 + \mu_0}{2}, \quad \lambda \text{ is the threshold for which } P_F = \alpha$$

$$P_F = P(Y > y_\lambda | H_0) = 1 - \Phi\left(\frac{y_\lambda - \mu_0}{\sigma}\right) = \alpha$$

$$P_D = P(Y > y_\lambda | H_1) = 1 - \Phi\left(\frac{y_\lambda - \mu_1}{\sigma}\right) \geq P_F = \alpha$$

Further simplify: solve $P_F = \alpha \Rightarrow \frac{y_\lambda - \mu_0}{\sigma} = \Phi^{-1}(1-\alpha)$

$$P_D = 1 - \Phi\left(\Phi^{-1}(1-\alpha) - \underbrace{\frac{\mu_1 - \mu_0}{\sigma}}_{\gamma \sim \text{SNR}}\right)$$



Plot α vs $P_D(\alpha)$ in MATLAB using $normcdf(\cdot)$ and $norminv(\cdot)$ functions

check:

$$\frac{\partial P_D}{\partial P_F} = \frac{\partial P_D / \partial y_\lambda}{\partial P_F / \partial y_\lambda} = L(y_\lambda) = \lambda = \text{slope at any given point}$$

Composite Hypothesis Testing

In general, $H_0 : p(\vec{y}|H_0, \theta)$, $H_1 : p(\vec{y}|H_1, \theta)$, where we know either

1. $p(\theta|H_i)$ for $i = 0$ and 1, or
2. θ is deterministic and $\theta \in \Lambda_i$ for $i = 0$ and 1.

Note: Simple hypothesis testing is a special case of composite testing, where θ is known under H_i.

Case (i): We know $p(\theta|H_i)$ and $p(\vec{y}|H_i, \theta)$. Then,

$$p(\vec{y}|H_i) = \int p(\theta|H_i) p(\vec{y}|H_i, \theta) \, d\theta \quad (*)$$

Now find the likelihood ratio

Example: $H_0 : Y \sim \mathcal{N}(0,1)$, $H_1 : Y \sim \mathcal{N}(\Theta,1)$,

$$p(\theta) = \begin{cases} e^{-\theta}, & \theta \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

From $(*)$,

$$\begin{aligned} p(y|H_1) &= \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y-\theta)^2}{2}\right) \exp(-\theta) d\theta \\ &= e^{-y+\frac{1}{2}} \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y-\theta-1)^2\right) d\theta \\ &= e^{-y+\frac{1}{2}} \left[1 - \Phi\left(\frac{0-(y-1)}{1}\right)\right] \\ &= e^{-y+\frac{1}{2}}\ \Phi(y-1) \end{aligned}$$

$\Rightarrow L(y) = \frac{p(y|H_1)}{p(y|H_0)}$

Case (ii) Consider the scenario in which, under H_i, $\vec{Y} \sim p(\vec{y}|H_i, \theta)$, $\theta \in \Lambda_i$

- optimal $P_D - P_F$ trade-off: Consider the Neyman-Pearson decision rule:

$$\max P_D$$
$$s.t.\ P_F \leq \alpha$$

Example: $H_0 : Y \sim \mathcal{N}(0, 1)$, $H_1 : Y \sim \mathcal{N}(\theta, 1)$, $\theta \in (0, \infty)$ is non-random and unknown

(Composite hypothesis testing with $\Lambda_1 = (0, \infty)$, $\Lambda_0 = 0$)

NP-test: given $\theta > 0$,

$$L_\theta(y) = \exp(\theta y - \frac{1}{2}\theta^2) \underset{0}{\overset{1}{\gtrless}} \lambda$$

$$\Rightarrow \theta y \underset{0}{\overset{1}{\gtrless}} \frac{1}{2}\theta^2 + \ln\lambda \Rightarrow y \underset{0}{\overset{1}{\gtrless}} \underbrace{\frac{1}{2}\theta + \frac{\ln\lambda}{\theta}}_{y_\lambda(\theta)}$$

To choose λ we solve $P_F = \alpha$

$$P_F = \int_{y_\lambda(\theta)}^{\infty} p(y|H_0)\,dy = \underbrace{1 - \Phi(y_\lambda(\theta)) = \alpha}_{y_\lambda(\theta)=\Phi^{-1}(1-\alpha)}$$

$$y \underset{0}{\overset{1}{\gtrless}} \Phi^{-1}(1-\alpha) \text{ for any given } \theta!$$

- In general, if the NP-optimal test is not a function of θ, the unknown parameter, the test is called "uniformly most powerful (UMP)". UMP tests are just as good as NP tests for which we know θ perfectly.

Example: $H_0 : Y \sim \mathcal{N}(0,1)$, $H_1 : Y \sim \mathcal{N}(\theta,1)$, $\theta \in \mathbb{R} \backslash \{0\}$. The NP test is UMP?

For any given θ

$$L_\theta(y) = \exp(\theta y - \frac{1}{2}\theta^2) \underset{0}{\overset{1}{\gtrless}} \lambda \Rightarrow \theta y \underset{0}{\overset{1}{\gtrless}} \ln \lambda + \frac{1}{2}\theta^2$$

- If $\theta > 0$, $\quad y \underset{0}{\overset{1}{\gtrless}} \frac{1}{2}\theta + \frac{1}{\theta} \ln \lambda$;
- If $\theta < 0$, $\quad y \underset{0}{\overset{1}{\gtrless}} \frac{1}{2}\theta + \frac{1}{\theta} \ln \lambda$

  where $y_\lambda = \frac{1}{2}\theta + \frac{1}{\theta} \ln \lambda$

- If $\theta > 0$, $\quad \Rightarrow \alpha = P_F = 1 - \Phi(y_\lambda) \; \Rightarrow \; y_\lambda = \Phi^{-1}(1-\alpha)$
- If $\theta < 0$, $\quad \Rightarrow \alpha = P_F = \int_{-\infty}^{y_\lambda} p(y|H_0)dy = \Phi(y_\lambda) \; \Rightarrow \; y_\lambda = \Phi^{-1}(\alpha)$

$$\delta_{NP}(y) = \begin{cases} 1, & \{y > \Phi^{-1}(1-\alpha), \theta > 0\} \; or \; \{y > \Phi^{-1}(\alpha), \theta < 0\} \\ 0, & \text{otherwise} \end{cases}$$

decision rule depends on $\theta \Rightarrow$ not UMP (check out suboptimal decision rules).

Locally Most Powerful (LMP) Test

The idea of LMP test is to design a test to perform well for a certain $\theta = \theta_0$.

Example: Under $H_i : \vec{Y} \sim g(\vec{y}|\theta \in \Lambda_i)$, $\Lambda_0 = \{\theta_0\}$ and $\Lambda_1 = (\theta_0, \infty)$

- We might want to optimize for $\theta \approx \theta_0$, where it is the most difficult to distinguish between H_0 and H_1.

Assumption:

Suppose for a given δ, P_D is differentiable fn. of θ around $\theta = \theta_0$.

$$
\begin{aligned}
P_D(\delta, \theta) &= P_D(\delta, \theta_0) + (\theta - \theta_0) P_D^{'}(\delta, \theta_0) + \text{ higher order terms} \\
&= P_F(\delta, \theta_0) = \alpha
\end{aligned}
$$

- For $\theta \approx \theta_0$,

$$\begin{array}{c} \max_{\delta} P_D(\delta,\theta) \\ s.t. \quad P_F(\delta,\theta) \le \alpha \end{array} \equiv \begin{array}{c} \max_{\delta} P_D^{'}(\delta,\theta_0) \\ s.t. \quad P_F(\delta,\theta_0) \le \alpha \end{array}$$

Lagrange multiplier approach:
$J(\delta,\theta_0,\lambda) = 1 - P_D^{'}(\delta,\theta_0) + \lambda(P_F(\delta,\theta_0) - \alpha)$

Solve: $\max_{\lambda>0} \min_{\delta} J(\delta,\theta_0,\lambda)$

- $P_D'(\delta,\theta_0) = \frac{\partial}{\partial\theta}\left[\int_{\Gamma_1} g(\vec{y}|\theta)\, d\vec{y}\right]\Big|_{\theta=\theta_0} = \int_{\Gamma_1} g'(\vec{y}|\theta_0)\, d\vec{y}$

relaxed problem:

$$\max_{\lambda>0}\left[(1-\lambda\alpha) + \min_{\delta}\int_{\Gamma_1}\left(\lambda\, g(\vec{y}|\theta_0) - g^{'}(\vec{y}|\theta_0)\right) d\vec{y}\right]$$

$$\Rightarrow \lambda\, g(\vec{y}|\theta_0) - g^{'}(\vec{y}|\theta_0) \underset{1}{\overset{0}{\gtrless}} 0$$

$$\Rightarrow \frac{g^{'}(\vec{y}|\theta_0)}{g(\vec{y}|\theta_0)} \underset{0}{\overset{1}{\gtrless}} \lambda, \quad \text{choose } \lambda \text{ to satisfy } P_F\Big|_{\theta=\theta_0} = \alpha$$

Notes:

1. if UMP exists, UMP $\equiv$ LMP

2. $$\frac{g(\vec{y}|\theta)}{g(\vec{y}|\theta_0)} = \frac{\partial}{\partial\theta}\left[\frac{g(\vec{y}|\theta)}{g(\vec{y}|\theta_0)}\right]\Bigg|_{\theta=\theta_0} = L'(\vec{y},\theta_0)$$

3. $$\frac{\partial}{\partial\theta}\underbrace{\log L(\vec{y},\theta)}_{LLR}\Big|_{\theta=\theta_0} = \frac{L'(\vec{y},\theta_0)}{L(\vec{y},\theta_0)} = L'(\vec{y},\theta_0)$$

Example: $g(y|\theta) = \frac{1}{\pi(1+(y-\theta)^2)}$

$H_i : Y \sim g(y|\theta \in \Lambda_i)$, $\Lambda_0 = \{0\}$, $\Lambda_1 = (0, \infty) \;\Rightarrow\; \theta_0 = 0$

$$L(y, \theta) = \frac{1 + y^2}{1 + (y - \theta)^2} \;\Rightarrow\; L'(y, 0) = \frac{2y}{1 + y^2}$$

LMP test: $\frac{2y}{1+y^2} \underset{0}{\overset{1}{\gtrless}} \lambda$ for λ s.t. $P_F\Big|_{\theta=0} = \alpha \;\rightarrow\;$ exercise

(can be shown that LMP test is a function of $\theta \Rightarrow$ UMP does not exist)

Signal with Unknown Parameters in Independent Noise:

Example: We study the case:

$H_0 : Y_k = N_k$, $H_1 : Y_k = N_k + \theta s_k$, $\theta > 0$, $\{N_k\}$ are independent, $1 \leq k \leq n$

Given θ, $L(\vec{y}, \theta) = \prod_{k=1}^{n} \frac{P_{N_k}(y_k - \theta s_k)}{P_{N_k}(y_k)}$

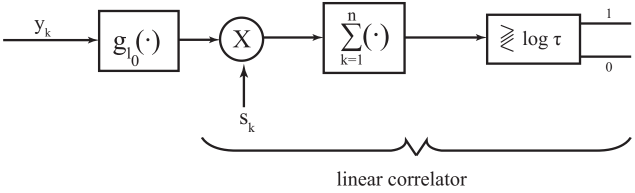
- LRTs depend on θ (unless $N_k \sim \mathcal{N}$) $\Rightarrow$ no UMP
- try LMP test around $\theta = 0$:

$$\Rightarrow \frac{\partial}{\partial \theta} \log L(\vec{y}, \theta)\Big|_{\theta=0} = \sum_{k=1}^{n} \frac{\partial}{\partial \theta} \left[\log P_{N_k}(y_k - \theta s_k) - \log P_{N_k}(y_k)\right] \Big|_{\theta=0}$$

$$= \sum_{k=1}^{n} s_k \underbrace{\left(-\frac{P'_{N_k}(y_k)}{P_{N_k}(y_k)}\right)}_{g_{l_0}(y_k)}$$

LMP test involves non-linear processing of y_k:

detector:



- N_k : Gaussian $\Rightarrow$ $g_{l_0}(y_k) = \frac{y_k}{\sigma^2}$ $\rightarrow$ linear correlator (makes sense, since UMP $\equiv$ LMP)
- N_k : Cauchy $\Rightarrow$ $P_{N_k}(y_k) = \frac{1}{\pi(1+y_k^2)}$ $\Rightarrow$ $g_{l_0}(y_k) = \frac{2y_k}{1+y_k^2}$

Generalized Likelihood Ratio Test (GLRT)

If UMP does not exist and LMP is not feasible (e.g., if g is not differentiable or if θ_0 is difficult to chose

$$L(\vec{y}) = \frac{\max\limits_{\theta \in \Lambda_1} p(\vec{y}|H_1, \theta)}{\max\limits_{\theta \in \Lambda_0} p(\vec{y}|H_0, \theta)} \underset{0}{\overset{1}{\gtrless}} \lambda, \ s.t. \ \underbrace{P_F\Big|_{\theta=\theta_0}}_{\arg\max\limits_{\theta \in \Lambda_0} p(\vec{y}|H_0,\theta)} = \alpha$$

Notes:

1. $\arg\max\limits_{\theta \in \Lambda_i} p(\vec{y}|H_i, \theta) \triangleq \hat{\theta}_i(\vec{y})$ is called the maximum likelihood estimate of θ under H_i given $\vec{y}$ is observed
2. GLRT does not satisfy any optimality criteria, but it works well in practice

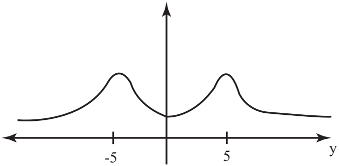
Example: $H_0 : Y \sim \mathcal{N}(0,1)$, $H_1 : Y \sim \mathcal{N}(\theta,1)$, $\theta \in \{-5,5\} \rightarrow$ no UMP test and for LMP, no clear choice for θ_0.

$\Rightarrow$ GLRT:

$$\hat{\theta}_1(y) = \arg\max_{\theta \in \{-5,5\}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y-\theta)^2\right) = \begin{cases} 5, & y > 0 \\ -5, & y \leq 0 \end{cases}$$

$$\Rightarrow [y - \hat{\theta}_1(y)]^2 = \begin{cases} (y-5)^2, & y > 0 \\ (y+5)^2, & y \leq 0 \end{cases} = (5 - |y|)^2$$

$$\Rightarrow p(y|H_1, \hat{\theta}_1(y)) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(5-|y|)^2\right)$$



GLRT:

$$\frac{\exp\left(-\frac{1}{2}(5-|y|)^2\right)}{\exp\left(-\frac{1}{2}y^2\right)} \underset{0}{\overset{1}{\gtrless}} \lambda, \quad \lambda \text{ satisfies } P_F = \alpha$$

Deterministic Signal in Colored Gaussian Noise

Noise: $\vec{N} \sim \mathcal{N}(\vec{0}, \Sigma_N) \Rightarrow P_{\vec{N}}(\vec{y}) = \frac{1}{(2\pi)^{n/2}[det(\Sigma_N)]^{1/2}} \exp(-\frac{1}{2}\vec{y}^T \Sigma_N^{-1} \vec{y})$.

Consider under: $H_i : \vec{Y} = \vec{s}_i + \vec{N}$

$$L(\vec{y}) = \frac{P_{\vec{N}}(\vec{y} - \vec{s}_1)}{P_{\vec{N}}(\vec{y} - \vec{s}_0)} = \exp\left((\vec{s}_1 - \vec{s}_0)^T \Sigma_N^{-1} (\vec{y} - \frac{\vec{s}_1 - \vec{s}_0}{2}) \right)$$

$$LLR(\vec{y}) = (\vec{s}_1 - \vec{s}_0)^T \Sigma_N^{-1} \vec{y} - \frac{1}{2} (\vec{s}_1 - \vec{s}_0)^T \Sigma_N^{-1} (\vec{s}_1 + \vec{s}_0) \underset{0}{\overset{1}{\gtrless}} \log \tau$$

$$(\vec{s}_1 - \vec{s}_0)^T \Sigma_N^{-1} \vec{y} \underset{0}{\overset{1}{\gtrless}} \log \tau + \frac{1}{2} (\vec{s}_1 - \vec{s}_0)^T \Sigma_N^{-1} (\vec{s}_1 + \vec{s}_0)$$

define: $T(\vec{Y}) = (\vec{s}_1 - \vec{s}_0)^T \Sigma_N^{-1} \vec{Y}$: Normal r.v. given H_i

$$\text{mean:} \quad E[T(\vec{Y})|H_i] = (\vec{s}_1 - \vec{s}_0)^T \Sigma_N^{-1} \vec{s}_i \;\rightarrow\; \mu_i$$

$$\begin{aligned}
\text{variance:} \quad var(T(\vec{Y})|H_i) &= var((\vec{s}_1 - \vec{s}_0)^T \Sigma_N^{-1} \vec{y}|H_i) \\
&= var((\vec{s}_1 - \vec{s}_0)^T \Sigma_N^{-1} \vec{N}) \\
&= E[(\vec{s}_1 - \vec{s}_0)^T \Sigma_N^{-1} \vec{N}\vec{N}^T \Sigma_N^{-1} (\vec{s}_1 - \vec{s}_0)] \\
&= (\vec{s}_1 - \vec{s}_0)^T \Sigma_N^{-1} (\vec{s}_1 - \vec{s}_0) \;\rightarrow\; \sigma^2
\end{aligned}$$

Signal Design: still assuming $H_i : \vec{s}_i + \vec{N}$, $\vec{N} \sim \mathcal{N}(\vec{0}, \Sigma_N)$

- What if we could choose $\{\vec{s}_0, \vec{s}_1\}$? We would choose it to maximize detection performance.

- Check out P_D, P_F:

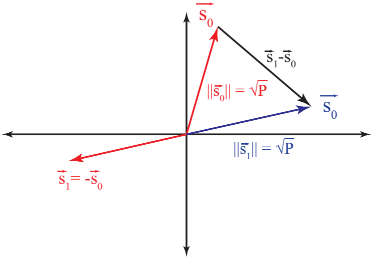
$$\delta(\vec{y}) = \begin{cases} 1, & T(\vec{y}) \geq y_\tau \\ 0, & T(\vec{y}) < y_\tau \end{cases}$$

$$P_D = 1 - \Phi\left(\frac{y_\tau - \mu_1}{\sigma}\right)$$

$$P_F = 1 - \Phi\left(\frac{y_\tau - \mu_0}{\sigma}\right)$$

- Maximize the difference: $\mu_1 - \mu_0 = (\vec{s}_1 - \vec{s}_0)^T \Sigma_N^{-1} (\vec{s}_1 + \vec{s}_0)$
- Choose $||\vec{s}_1 - \vec{s}_0||^2 = \infty \;\Rightarrow\; P_D = 1, P_F = 0 \;\rightarrow$ impractical

Power constraint: $||\vec{s}_i||^2 \sim$ energy of the signal $= \sum_{k=1}^{n} s_{i,k}^2 \rightarrow ||\vec{s}_i||^2 \leq P$



alternate formulation: Let $\vec{s}_1 - \vec{s}_0 \triangleq \vec{s}$

$$\max_{\vec{s} \in \mathbb{R}^n} \quad \vec{s}^T \Sigma_N^{-1} \vec{s}$$

$$s.t. \quad ||\vec{s}|| \leq 2\sqrt{P} \quad \text{and} \quad ||\vec{s}_1|| \leq \sqrt{P}, \; ||\vec{s}_\theta|| \leq \sqrt{P}$$

$$\max_{\vec{s}\in\mathbb{R}^n} \quad \vec{s}^T \Sigma_N^{-1} \vec{s}$$

$$s.t. \quad ||\vec{s}|| \leq 2\sqrt{P} \text{ and } ||\vec{s_1}|| \leq \sqrt{P}, \quad ||\vec{s}_0|| \leq \sqrt{P}$$

objective: $\vec{s}^T \Sigma_N^{-1} \vec{s} = \vec{s}^T Q \Lambda_N^{-1} Q^T \vec{s}$

$\Sigma_N = Q \Lambda_N Q^T$

where $Q = [\vec{q_1}, \vec{q_2}, ..., \vec{q_n}]$, and $\Lambda_N = diag(\lambda_1, \lambda_2, ..., \lambda_n)$

$$\vec{s}_Q \triangleq \vec{s}^T Q = [\langle \vec{s}, \vec{q_1} \rangle, \langle \vec{s}, \vec{q_2} \rangle, ..., \langle \vec{s}, \vec{q_n} \rangle]$$

$$\vec{s}_Q \Lambda_N^{-1} = \left[\frac{\langle \vec{s}, \vec{q_1} \rangle}{\lambda_1}, \frac{\langle \vec{s}, \vec{q_2} \rangle}{\lambda_2}, ..., \frac{\langle \vec{s}, \vec{q_n} \rangle}{\lambda_n} \right]$$

$$\Rightarrow \underbrace{\vec{s}^T \Sigma_N^{-1} \vec{s}}_{maximize} = \langle \vec{s}_Q \Lambda_N^{-1}, \vec{s}_Q \rangle = \sum_{i=1}^{n} \frac{1}{\lambda_i} \langle \vec{q}_i, \vec{s} \rangle^2$$

$$\leq \frac{1}{\lambda_{min}} \sum_{i=1}^{n} \langle \vec{q}_i, \vec{s} \rangle^2 \quad (*)$$

$$= \frac{1}{\lambda_{min}} ||\vec{s}||^2$$

What if we choose $\vec{s} = 2\sqrt{P}\vec{q}_{min}$ (q_{min} is the eigenvector associated with λ_{min})

$\Rightarrow$ $(*)$ is satisfied with equality

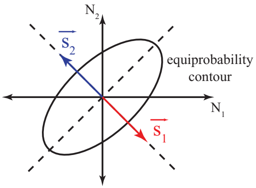
$\Rightarrow$ choose $\vec{s}_0$ and $\vec{s}_1$ along $\vec{q}_{min} \rightarrow \vec{s}_0 = \sqrt{P}\ \vec{q}_{min},\ \vec{s}_1 = -\sqrt{P}\ \vec{q}_{min}$ (or vice versa)

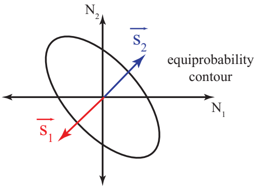
and therefore $\vec{s}^T \Sigma_N^{-1} \vec{s} = \frac{4P}{\lambda_{min}}$

Example: Let $\Sigma_N = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$, $|\rho| < 1$

$$\Rightarrow \lambda_{1,2} = 1 \mp \rho \ \rightarrow \ Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

- If $\rho > 0$, $\lambda_{min} = 1 - \rho$ $\Rightarrow \vec{q}_{min} = \frac{1}{2}[1, -1]^T \ \Rightarrow \ \vec{s}_i = \mp\sqrt{\frac{P}{2}}[1, -1]^T$
- If $\rho < 0$, $\lambda_{min} = 1 + \rho$, $\vec{q}_{min} = \frac{1}{\sqrt{2}}[1, 1]^T \ \Rightarrow \ \vec{s}_i = \mp\sqrt{\frac{P}{2}}[1, 1]^T$





Back to detection under colored Gaussian noise:

Whitening Techniques

Q: $\vec{Y}$ is colored Gaussian, can I find an A such that, $A\vec{Y}$ is uncorrelated (white) and use simple correlator for detection?

recall that: under $H_i \rightarrow \vec{Y} = \vec{s}_i + \vec{N}$, where $|\Sigma_N| > 0$

consider: $\vec{Y}_t = A\vec{Y} = \Lambda_N^{-1/2} Q^T \vec{Y} = \vec{s}_{i_t} + \vec{N}_t$,

where: $\vec{s}_{i_t} = \Lambda_N^{-1/2} Q^T \vec{s}_i$, $\vec{N}_t = \Lambda_N^{-1/2} Q^T \vec{N}$.

Check:

$$
\begin{aligned}
E[\vec{N}_t] &= \Lambda_N^{-1/2} Q^T E[\vec{N}] = \vec{0} \\
\Sigma_{N_t} &= E[\vec{N}_t \vec{N}_t^T] = E[\Lambda_N^{-1/2} Q^T \vec{N} \vec{N}^T Q \Lambda_N^{-1/2}] \\
&= \Lambda_N^{-1/2} Q^T \underbrace{\Sigma_N}_{Q\Lambda_N Q^T} Q \Lambda_N^{-1/2} = \Lambda_N^{-1/2} \Lambda_N \Lambda_N^{-1/2} \\
&= \mathbf{I} \Rightarrow \text{noise samples iid } \mathcal{N}(0,1) \text{ after transformation}
\end{aligned}
$$

- Since A is an invertible matrix, the problem can be transformed back to the original. Thus, no loss of information after transformation.