

Final Exam
Saturday, May 7, 2005

Name SOLUTION SET

Score _____

Please do not turn this page until requested to do so.

This is a closed-book exam; you may however use three cheat sheets. Make sure to show all of your work on the longer problems for partial credit. Box your answers, and write neatly. Backs of pages may be used for scratch work if necessary. No calculators are allowed.

Problem 1 (20 pts). Consider a point s on the unit circle in \mathbb{R}^2 . You are given n iid points

$$Y_i = s + N_i, \quad 1 \leq i \leq n,$$

where each $N_i \in \mathbb{R}^2$ is Gaussian with mean zero and covariance matrix $C = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, where $|\rho| < 1$.

(a) (15 pts) Determine the ML estimator of s .

Equivalently, write $s = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$ and derive $\hat{\theta}_{ML}$.

$$\text{We have } \hat{\theta}_{ML} = \operatorname{argmax}_{0 \leq \theta < 2\pi} \ln P(Y|\theta)$$

$$= \operatorname{argmin}_{0 \leq \theta < 2\pi} \sum_{i=1}^n (Y_i - s(\theta))^T C^{-1} (Y_i - s(\theta))$$

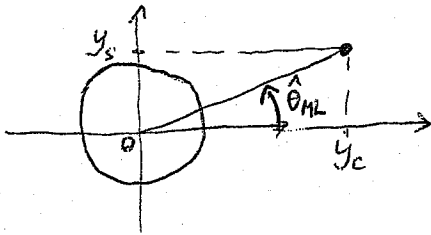
$$\rightarrow = \frac{1}{1-\rho^2} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}$$

$$= \operatorname{argmax}_{0 \leq \theta < 2\pi} \left(\sum_{i=1}^n Y_i \right)^T \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix} \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$$

$$= \operatorname{argmax}_{0 \leq \theta < 2\pi} (y_c \cos \theta + y_s \sin \theta)$$

$$\text{where } y_c = \sum_{i=1}^n (y_{i1} - \rho y_{i2})$$

$$y_s = \sum_{i=1}^n (y_{i2} - \rho y_{i1})$$



$$\boxed{\hat{\theta}_{ML} = \arctan \frac{y_s}{y_c}}$$

(b) (5 pts) Is this estimator a sufficient statistic for estimating s ? A minimal sufficient statistic? Explain your answer.

$$\text{Can write } P(Y|\theta) = h(y) e^{y_c \cos \theta + y_s \sin \theta} = h(y) e^{\sqrt{y_c^2 + y_s^2} \cos(\theta - \hat{\theta}_{ML})}$$

From Factorization Theorem, (y_c, y_s) is a sufficient statistic

for estimating θ but $\hat{\theta}_{ML} = \arctan \frac{y_s}{y_c}$ is not!

Problem 2 (20 pts). Consider the hypothesis test

$$\begin{cases} H_0 : Y_i = N_i, & 1 \leq i \leq n \\ H_1 : Y_i = s + N_i, & 1 \leq i \leq n \end{cases}$$

where s is on the unit circle in \mathbb{R}^2 (you don't know where), and the vectors N_i are iid normal.

(a) (10 pts) Is there a Uniformly Most Powerful test for this problem? Explain why. \rightarrow Take $s=0$ in Pbm 1.

As in Pbm 1, write $s = s(\theta) = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$

For known θ , NP rule is of the form $y_c \cos \theta + y_s \sin \theta \underset{H_0}{\overset{H_1}{\geq}} \tau$

clearly depends on θ

\Rightarrow no UMP test

(b) (10 pts) Derive the Generalized Likelihood Ratio Test for this problem, and sketch the boundary of the decision region.

From Pbm 1a we have $\hat{\theta}_{ML} = \arctan \frac{y_s}{y_c}$ where $\begin{cases} y_c = \sum_{i=1}^n y_{c1} \\ y_s = \sum_{i=1}^n y_{c2} \end{cases}$
(with $s=0$)

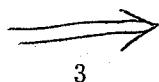
GLRT is of the form

$$l_G(y) = \frac{1}{2} \sum_{i=1}^n \left[\|y_i\|^2 - \|y_i - s(\hat{\theta})\|^2 \right] \underset{H_0}{\overset{H_1}{\geq}} \ln \tau$$

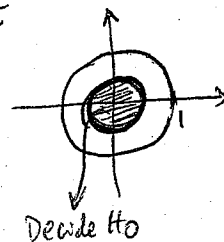
$$s(\hat{\theta})^T \left(\sum_{i=1}^n y_i \right) - \frac{n}{2} \geq \ln \tau$$

$$y_c \cos \hat{\theta} + y_s \sin \hat{\theta} + \frac{n}{2} \geq \ln \tau$$

$$R \stackrel{\Delta}{=} \frac{1}{n} \sqrt{\left(\sum_i y_{c1} \right)^2 + \left(\sum_i y_{c2} \right)^2}$$



$$R \underset{H_0}{\overset{H_1}{\geq}} \frac{1}{2} + \frac{\ln \tau}{n}$$



Problem 3 (20 pts) Consider the binary hypothesis testing problem

$$\begin{cases} H_0 : Y_k = N_k, & 1 \leq k \leq 10 \\ H_1 : Y_k = S_k + N_k, & 1 \leq k \leq 10 \end{cases}$$

where N_k are independent Gaussian noise samples with mean zero and variances k .

(a) (10 pts) Design the signal $S_k, 1 \leq k \leq 10$ so as to minimize probability of error, subject to the energy constraint $\sum_{k=1}^{10} S_k^2 \leq E$.

$$\min_{\mathbf{s}} P_e \Leftrightarrow \max_{\mathbf{s}} \left(d^2 = \sum_{k=1}^{10} \frac{S_k^2}{k} \right) \quad \text{subj. to } \sum_{k=1}^{10} S_k^2 \leq E$$

$$\Rightarrow \boxed{\begin{matrix} S_1 = \pm \sqrt{E} \\ S_k = 0 \quad \forall k > 1 \end{matrix}} \Rightarrow \boxed{d^2 = E}$$

(b) (10 pts) It is known that one sample from the sequence $Y_k, 1 \leq k \leq 10$ might be erased, so the detector might have access to only 9 samples. The location of any missing sample is known to the detector, but unknown to the signal designer. Repeat the optimal signal design problem of Part (a) to minimize the maximum (over all possible erasures) probability of error.

Hint. There are two equally good solutions.

$$\min_{\mathbf{s}} P_e \Leftrightarrow \max_{\mathbf{s}} \left(d^2 = \min_{1 \leq i \leq 10} \sum_{k \neq i} \frac{S_k^2}{k} \right) \quad \text{subj. to } \sum_{k=1}^{10} S_k^2 \leq E$$

Consider spreading the energy over the first \tilde{K} components, and try "equalizer design": $\frac{S_k^2}{k} = C$ for all $1 \leq k \leq \tilde{K} \Rightarrow d^2 = C(\tilde{K}-1)$

To satisfy energy constraint, need $E \geq \sum_{k=1}^{\tilde{K}} S_k^2 = C \sum_{k=1}^{\tilde{K}} k = \frac{C \tilde{K}(\tilde{K}+1)}{2}$

clearly want equality

$$\left. \begin{aligned} d^2 &= c(\tilde{k}-1) \\ E &= \frac{c\tilde{k}(\tilde{k}+1)}{2} \end{aligned} \right\} \Rightarrow d^2 = E \cdot \frac{2(\tilde{k}-1)}{\tilde{k}(\tilde{k}+1)}$$

$$\tilde{k} = 1 \Rightarrow d^2 = 0 \quad (\text{obviously})$$

$$\tilde{k} = 2 \Rightarrow d^2 = \frac{E}{3}$$

$$\tilde{k} = 3 \Rightarrow d^2 = \frac{E}{3}$$

$$\tilde{k} = 4 \Rightarrow d^2 = \frac{3}{10} E$$

\vdots
 d^2 decreasing in \tilde{k} for $\tilde{k} > 3$.

\Rightarrow Can achieve $d^2 = \frac{E}{3}$ using $s_k = \pm \sqrt{k} c$ for $1 \leq k \leq \tilde{k}$

$$\text{Design \#1} : \tilde{k} = 2 \Rightarrow c = \frac{2E}{\tilde{k}(\tilde{k}+1)} = \frac{E}{3} \Rightarrow \boxed{s = \left(\pm \sqrt{\frac{E}{3}}, \pm \sqrt{\frac{2E}{3}} \right)}$$

$$\text{Design \#2} : \tilde{k} = 3 \Rightarrow c = \frac{2E}{\tilde{k}(\tilde{k}+1)} = \frac{E}{6} \Rightarrow \boxed{s = \left(\pm \sqrt{\frac{E}{6}}, \pm \sqrt{\frac{2E}{6}}, \pm \sqrt{\frac{3E}{6}} \right)}$$

Nonequalizer designs are clearly suboptimal:

Could improve $\frac{s_k^2}{k}$ by slightly decreasing $\frac{s_l^2}{l}$ for some other $l \neq k$.
 (for worst k)

Problem 4 (20 pts) The random variable W is equal to 0 with probability $1 - \epsilon$, and to a $\mathcal{N}(0, 1)$ random variable with probability ϵ . Given a length- n iid sequence \mathbf{W} generated from this distribution and a known signal \mathbf{s} , consider the binary hypothesis test

$$\begin{cases} H_0 : \mathbf{Y} = -\mathbf{s} + \mathbf{W} \\ H_1 : \mathbf{Y} = \mathbf{s} + \mathbf{W} \end{cases}$$

Derive an upper bound on the probability of error for that test; the bound should be tight in the exponent as $n \rightarrow \infty$.

Hint: Recall the definition of the Dirac impulse: $\int f(x)\delta(x) dx = \lim_{\epsilon \rightarrow 0} \frac{1}{2\epsilon} \int_{-\epsilon}^{\epsilon} f(x) dx$.

$$\begin{cases} P_{0k}(y_k) = (1-\epsilon) \delta(y_k + s_k) + \epsilon \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_k + s_k)^2} \\ P_{1k}(y_k) = (1-\epsilon) \delta(y_k - s_k) + \epsilon \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_k - s_k)^2} \end{cases}$$

Chernoff integral: $C_s(P_{0k}, P_{1k}) = \int P_{0k}^{1-s}(y) P_{1k}^s(y) dy$ $0 < s < 1$
for component k

$$= \epsilon C_s(\mathcal{N}(-s_k, 1), \mathcal{N}(s_k, 1))$$

$$= \epsilon e^{-\frac{(2s_k)^2 s(1-s)}{2}}$$

$$\text{Chernoff distance } d_s(P_0, P_1) = - \sum_{k=1}^n \ln C_s(P_{0k}, P_{1k})$$

$$= -n \ln \epsilon + 2s(1-s) \|\mathbf{s}\|^2$$

is maximized by $s = \frac{1}{2}$ (Bhattacharyya)

$$\Rightarrow b(P_0, P_1) = d_{\frac{1}{2}}(P_0, P_1) = -n \ln \epsilon + \frac{1}{2} \|\mathbf{s}\|^2$$

$$\Rightarrow P_e \leq e^{-b(P_0, P_1)} = \epsilon^n e^{-\frac{1}{2} \|\mathbf{s}\|^2}$$

tight in the exponent

Problem 5 (20 pts) Let $Y_i = \theta + W_i$, $1 \leq i \leq n$, where W_i are iid Laplacian random variables with variance σ^2 .

$$P_W(w) = \frac{1}{\sqrt{2}\sigma} e^{-\sqrt{2}|w|/\sigma}$$

(a) (15 pts) Prove that for any constant $c > 0$ we have

$$\lim_{n \rightarrow \infty} \Pr \left[|\text{med}\{Y_1, \dots, Y_n\} - \theta| \geq n^{-1/2}c \right] \leq \exp \left\{ -\frac{c^2}{2\sigma^2} \right\}$$

where $\text{med}\{\dots\}$ denote the median value of a set of numbers.

Observe that $\text{med}\{Y_1, \dots, Y_n\} = \hat{\theta}_{ML}(Y)$.

The estimation errors for MLE are asymptotically Gaussian:

$$(*) \quad \sqrt{n} (\hat{\theta}_{ML} - \theta) \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{I_\theta}\right)$$

$$\text{where } I_\theta = \int_{-\infty}^{\infty} \frac{(P'_W(w))^2}{P_W(w)} dw \quad (\text{location parameter } \theta)$$

$$= E \left[\left(\ln P'_W(w) \right)^2 \right]$$

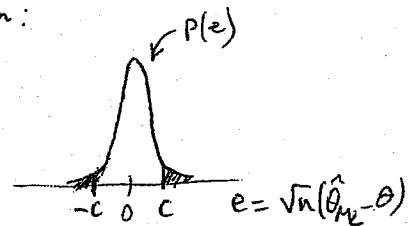
$$= E \left(\frac{z^2}{\sigma^2} \right) = \frac{2}{\sigma^2}$$

$$\Rightarrow \text{asymptotic variance of MLE} = \frac{\sigma^2}{2n}$$

$$\text{Due to } (*), \text{ we have } \Pr \left[|\hat{\theta}_{ML} - \theta| \geq n^{-1/2}c \right] = \Pr \left[\sqrt{n} |\hat{\theta}_{ML} - \theta| \geq c \right]$$

$$\approx 2Q \left(\frac{c}{\sigma/\sqrt{2}} \right)$$

$$\leq \exp \left\{ -\frac{c^2}{2\sigma^2} \right\}$$



Q.E.D.

(b) (5 pts). Is it also true that

$$\lim_{n \rightarrow \infty} \exp \left\{ \frac{nc^2}{2\sigma^2} \right\} \Pr \left[|\text{med}\{Y_1, \dots, Y_n\} - \theta| \geq c \right] \leq 1?$$

Explain why.

Not necessarily because $\Pr \left[|\hat{\theta}_{ML} - \theta| \geq c \right]$ can only be determined using a large-deviations analysis: the tail probabilities are non-Gaussian