Signal Detection + Estimation

4F6

2006

(Solutions)

2

Q 1

The first part is

book work

20%

3

1 cont).

For the scalar exponential family, the likelihood for the data $\underline{x}$ is given by

$$p(\underline{x} \mid \theta) = \prod_{n=0}^{N-1} \exp\left(A(\theta) B(x_n) + C(x_n) + D(\theta)\right)$$

$$= \exp\left(A(\theta) \sum_n B(x_n) + \sum_n C(x_n) + N D(\theta)\right)$$

$$= \exp\left(A(\theta) \sum_n B(x_n) + N D(\theta)\right) \times \exp\left(\sum_n C(x_n)\right).$$

$$\equiv g\left(T(\underline{x}), \theta\right) \times h(\underline{x})$$

Hence using the Neyman-Fisher factorization theorem, the sufficient statistic is given by

$$T(x) = \sum_n B(x_n).$$

40%.

1 cont.)

For the Gaussian case, we can write

$$p(x|\mu) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x-\mu)^2\right)$$

(assuming unit variance for simplicity – makes no difference

$$\therefore p(x|\mu) = \exp\Big(\underbrace{x\mu}_{A(\mu)B(x)} \underbrace{- \tfrac{1}{2}x^2}_{C(x)} + \underbrace{\left(-\tfrac{1}{2}\mu^2 + \ln 1/\sqrt{2\pi}\right)}_{D(\mu)}\Big)$$

$$\therefore \quad T(x) = \sum_{n=0}^{N-1} x_n$$

For the exponential case,

$$p(x|\lambda) = \lambda e^{-\lambda x}$$

$$= \exp \underbrace{-\lambda x}_{A(\lambda)B(x)} + \underbrace{0}_{C(x)} + \underbrace{\ln \lambda}_{D(\lambda)}$$

$$\therefore \quad T(x) = \sum_{n=0}^{N-1} x_n$$

40%

5/

2/ first part is book work. 35%.

——//——

parameterizing A, and B as $\theta = \begin{bmatrix} A \\ B \end{bmatrix}$,

the 2×2 Fisher information is

$$I(\theta) = \begin{bmatrix} -E \frac{\partial^2}{\partial A^2} \ln p(d|\theta) & -E \frac{\partial^2 \ln p(d|\theta)}{\partial A \partial B} \\ -E \frac{\partial^2 \ln p(d|\theta)}{\partial B \partial A} & -E \frac{\partial^2 \ln p(d|\theta)}{\partial B^2} \end{bmatrix}$$

and $p(d|\theta) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (d(n) - A - Bn)^2\right).$

$$\frac{\partial \ln p(d|\theta)}{\partial A} = \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (d(n) - A - Bn) \; ; \quad \frac{\partial^2 \ln p(d|\theta)}{\partial A^2} = -\frac{N}{\sigma^2}$$

$$\frac{\partial \ln p(d|\theta)}{\partial B} = \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (d(n) - A - Bn)\, n \; ; \quad \frac{\partial^2 \ln p(d|\theta)}{\partial B^2} = -\frac{1}{\sigma^2} \sum_{n=0}^{N-1} n^2$$

$$\frac{\partial \ln p(d|\theta)}{\partial A \partial B} = -\frac{1}{\sigma^2} \sum_{n=0}^{N-1} n$$

6/

2 cont

$$I(\theta) = \frac{1}{\sigma^2}\begin{bmatrix} N & \sum n \\ \sum n & \sum n^2 \end{bmatrix}.$$

$$= \frac{1}{\sigma^2}\begin{bmatrix} N & \frac{N(N-1)}{2} \\ \frac{N(N-1)}{2} & \frac{N(N-1)(2N-1)}{6} \end{bmatrix}.$$

30%

$$\therefore I^{-1} = \sigma^2 \begin{bmatrix} \frac{2(2N-1)}{N(N+1)} & -\frac{6}{N(N+1)} \\ -\frac{6}{N(N+1)} & \frac{12}{N(N^2-1)} \end{bmatrix}.$$

$$\therefore \operatorname{var}(\hat{A}) \geq \frac{2(2N-1)\sigma^2}{N(N+1)}$$

$$\operatorname{var}(\hat{B}) \geq \frac{12\sigma^2}{N(N^2-1)}$$

$$\therefore \frac{\operatorname{var}(\hat{A})}{\operatorname{var}(\hat{B})} = \frac{(2N-1)(N-1)}{6} > 1 \quad \text{for } N \geq 3.$$

$\therefore$ easier to estimate B.

35%

Q3 First part is book work. 40%

—//—

The general linear model may be written

$$\underline{d} = \underline{G}\,\underline{\theta} + \underline{w}$$

where the model parameters are contained in the vector $\underline{\theta}$. $\underline{d}$ is the observed data vector, $\underline{w}$ is the noise vector and $\underline{G}$ is a matrix.

For the signal

$$s(n) = A + Bn$$

the observed data is $\underline{d} = \underline{s} + \underline{w}$

$$\underline{d} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & N-1 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} + \underline{w} = \underline{G}\,\underline{\theta} + \underline{w}.$$

8/

3 cont

The likelihoods for the observed noisy data $d(n)$ are

$$p(d|H_0) = \frac{1}{(2\pi)^{N/2}|C|^{1/2}} e^{-\frac{1}{2}d^T C^{-1} d}.$$

$$p(d|H_1) = \frac{1}{(2\pi)^{N/2}|C|^{1/2}} e^{-\frac{1}{2}(d-s)^T C^{-1}(d-s)}.$$

and the NP detector is

$$L(d) = \frac{p(d|H_1)}{p(d|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda.$$

$$\therefore L(d) = e^{-\frac{1}{2}\left[-2d^T C^{-1} s + s^T C^{-1} s\right]}$$

$$\therefore d^T C^{-1} s \underset{H_0}{\overset{H_1}{\gtrless}} \frac{1}{2} s^T C^{-1} s + \log(\lambda).$$

$$\therefore \frac{1}{\sigma^2} d^T \underline{H}\,\underline{\theta} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda' \qquad \therefore \boxed{\frac{1}{\sigma^2}\sum d(n)(A+Bn) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda'}$$

60%

Q4 Note that these solutions are given with rather more detail and explanation than would be expected from candidates.

A reasonable criterion to use for deciding between two alternative hypotheses is:

$$P(H_1|\mathrm{y}) \underset{H_0}{\overset{H_1}{\gtrless}} P(H_0|\mathrm{y}) \tag{1}$$

Using Bayes' rule, the two posterior probabilities may be expressed in terms of likelihood functions and prior probabilities as follows:

$$P(H_0|\mathrm{y}) = \frac{P(\mathrm{y}|H_0)\,P(H_0)}{P(\mathrm{y})}$$

or

$$P(H_1|\mathrm{y}) = \frac{P(\mathrm{y}|H_1)\,P(H_1)}{P(\mathrm{y})}$$

$$P(H_0|\mathrm{y}) = \lim_{\mathrm{dy}\to 0} \frac{P(\mathrm{y} \le \mathrm{Y} < \mathrm{y}+\mathrm{dy}|H_0)\,P(H_0)}{P(\mathrm{y} \le \mathrm{Y} < \mathrm{y}+\mathrm{dy})}$$

$$\therefore\ P(H_0|\mathrm{y}) = \frac{p(\mathrm{y}|H_0)\,P(H_0)}{p(\mathrm{y})} \tag{2}$$

where $p(\mathrm{y}|H_0)$ is the likelihood function and $P(H_0)$ is the a-priori probability of hypothesis H_0.
Similarly:

$$P(H_1|\mathrm{y}) = \frac{p(\mathrm{y}|H_1)\,P(H_1)}{p(\mathrm{y})} \tag{3}$$

Substituting equations 2 and 3 in equation 1 gives:

$$p(\mathrm{y}|H_1)\,P(H_1) \underset{H_0}{\overset{H_1}{\gtrless}} p(\mathrm{y}|H_0)\,P(H_0)$$

Thus the vector form of the MAP criterion becomes:

$$\boxed{\frac{p(\mathrm{y}|H_1)}{p(\mathrm{y}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P(H_0)}{P(H_1)}}$$

[30%]

4 cont

Let $\mathbf{R}_0 \in \Re^N$ be the region where the values of $\mathbf{y}$ are such that:

$$\frac{p(\mathbf{y}|H_1)}{p(\mathbf{y}|H_0)} < \frac{P(H_0)}{P(H_1)}$$

i.e. the region where the decision is D_0 (accept hypothesis H_0).

Similarly, let $\mathbf{R}_1 \in \Re^N$ be the region where the values of $\mathbf{y}$ are such that:

$$\frac{p(\mathbf{y}|H_1)}{p(\mathbf{y}|H_0)} > \frac{P(H_0)}{P(H_1)}$$

i.e. the region where the decision is D_1 (accept hypothesis H_1).

The false alarm probability is:

$$P(D_1|H_0) = \int_{\mathbf{y}\in\mathbf{R}_1} p(\mathbf{y}|H_0)\, \mathrm{d}y$$

and the miss probability is:

$$P(D_0|H_1) = \int_{\mathbf{y}\in\mathbf{R}_0} p(\mathbf{y}|H_1)\, \mathrm{d}y$$

The average error probability is:

$$P_e = P(D_1|H_0)\, P(H_0) + P(D_0|H_1)\, P(H_1)$$

[20%]

The probability density function for a Gaussian process $\mathbf{d}$ is:

$$p(\mathbf{d}) = \frac{1}{(2\pi)^{\frac{N}{2}}\, |\mathbf{C}|^{\frac{1}{2}}} e^{-\frac{1}{2}\mathbf{d}^T C^{-1} d}$$

where the noise covariance matrix $\mathbf{C}$ is given by:

$$\mathbf{C} = E\{\mathbf{d}\,\mathbf{d}^{\mathrm{T}}\}$$

and

$$\mathbf{d} = \begin{bmatrix} d(1) \\ d(2) \\ \vdots \\ d(N) \end{bmatrix}$$

and $\mathbf{m}$ is the mean noise vector:

$$\mathbf{m} = E\{\mathbf{d}\}$$

4 cont

If the channel noise is white then the noise samples at the different time instants will be statistically independent and the noise covariance matrix will be diagonal with the noise variance σ^2 on the diagonal.

If the received signal samples $y(1)\ y(2)\ldots y(N)$ are formed into the vector $\mathbf{y}$ then the likelihood functions can be written as :

$$p(\mathbf{y}|H_0) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} e^{-\frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{s}_0)^T(\mathbf{y}-\mathbf{s}_0)} \tag{4}$$

$$p(\mathbf{y}|H_1) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} e^{-\frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{s}_1)^T(\mathbf{y}-\mathbf{s}_1)} \tag{5}$$

The MAP likelihood ratio test is:

$$L(\mathbf{y}) = \frac{p(\mathbf{y}|H_1)}{p(\mathbf{y}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P(H_0)}{P(H_1)}$$

Substituting for $p(\mathbf{y}|H_0)$ and $p(\mathbf{y}|H_1)$ from equations 4 and 5 gives:

$$L(\mathbf{y}) = e^{-\frac{1}{2\sigma^2}[-2\mathbf{y}^T\mathbf{s}_0+\mathbf{s}_0^T\mathbf{s}_0+2\mathbf{y}^T\mathbf{s}_1+\mathbf{s}_1^T\mathbf{s}_1]}$$

Taking logarithms the likelihood ratio test becomes:

$$-\frac{1}{2\sigma^2}[-2\mathbf{y}^T\mathbf{s}_0+\mathbf{s}_0^T\mathbf{s}_0+2\mathbf{y}^T\mathbf{s}_1+\mathbf{s}_1^T\mathbf{s}_1] \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P(H_0)}{P(H_1)}$$

and binary symbols 0 and 1 are equiprobable ie. $P(H_0) = P(H_1) = 0.5$

$$\therefore\ \mathbf{y}^T(\mathbf{s}_1-\mathbf{s}_0) \underset{H_0}{\overset{H_1}{\gtrless}} \frac{1}{2}(\mathbf{s}_1^T\mathbf{s}_1-\mathbf{s}_0^T\mathbf{s}_0)$$

[30%]

This is the Likelihood Ratio detector to discriminate between the two

hypotheses H_0 and H_1; i.e. that signal $\mathbf{s}_0$ is present or signal $\mathbf{s}_1$ is present. The vector dot product $\mathbf{y}^T(\mathbf{s}_1-\mathbf{s}_0)$ is compared with a threshold value $\frac{1}{2}(\mathbf{s}_1^T\mathbf{s}_1-\mathbf{s}_0^T\mathbf{s}_0)$.

When the channel noise is coloured (ie. correlated) the preceding theory may be redeveloped with the appropriate noise model. The results show that the MAP detector has the same form as above but the received signal is first transformed by a matrix which has the effect of decorrelating or whitening the noise. However the matrix also operates on the signal part ($\mathbf{s}_0$ or $\mathbf{s}_1$) so the signal terms in the MAP detector must be replaced with transformed signal terms. This matrix operation is often known as *prewhitening*.

[20%]