EUROCOM 07/02/2014
SSP Statistical Signal Processing

Exam

You can answer the questions in French or in English. If you get stuck in a particular part of a question, do not hesitate to try to answer the rest of the question or start the next question. Results that are available in the course notes can be quoted and do not need to be rederived.

Parameter Estimation

1. On the Beneficial Bias of MMSE Estimation

Consider the Bayesian linear model $Y = H\,\theta + V$ with $\theta \sim \mathcal{N}(0, C_{\theta\theta})$ and $V \sim \mathcal{N}(0, C_{VV})$ independent (we consider here $m_\theta = 0$ for simplicity).

(a) The LMMSE estimator is $\widehat{\theta}_{LMMSE} = C_{\theta Y} C_{YY}^{-1} Y = (C_{\theta\theta}^{-1} + H^T C_{VV}^{-1} H)^{-1} H^T C_{VV}^{-1} Y$.
What are the unconstrained MMSE and the MAP estimators?

(b) What is the error covariance matrix ?

$$R_{\widetilde{\theta}\widetilde{\theta}}^{LMMSE} = E_\theta E_{Y|\theta}\, \widetilde{\theta}_{LMMSE} \widetilde{\theta}_{LMMSE}^T \tag{1}$$

(c) The conditional bias of an estimator $\widehat{\theta}$ is $b_{\widehat{\theta}}(\theta) = E_{Y|\theta}\widehat{\theta}(Y) - \theta$.
The BLUE estimator is the LMMSE estimator under the constraint of conditional unbiasedness. So $b_{BLUE}(\theta) = 0$.
What is $\widehat{\theta}_{BLUE}$ in terms of the quantities appearing in the Bayesian linear model considered here?
Is there another classical deterministic estimator that equals $\widehat{\theta}_{BLUE}$ in this case?

(d) What is the error covariance matrix ?

$$R_{\widetilde{\theta}\widetilde{\theta}}^{BLUE} = E_\theta E_{Y|\theta}\, \widetilde{\theta}_{BLUE} \widetilde{\theta}_{BLUE}^T \tag{2}$$

(e) Show that $R_{\widetilde{\theta}\widetilde{\theta}}^{LMMSE} \leq R_{\widetilde{\theta}\widetilde{\theta}}^{BLUE}$.
Note that this is true in spite of $\widehat{\theta}_{LMMSE}$ being (conditionally) biased and $\widehat{\theta}_{BLUE}$ being unbiased.

(f) Returning to $\widehat{\theta}_{LMMSE}$, what is the bias $b_{LMMSE}(\theta)$?

(g) Show that

$$R_{\widetilde{\theta}\widetilde{\theta}} = \underbrace{E_\theta\, b_{\widehat{\theta}}(\theta)\, b_{\widehat{\theta}}^H(\theta)}_{\text{(bias)}^2} + \underbrace{E_\theta E_{Y|\theta}\, (\widehat{\theta} - E_{Y|\theta}\widehat{\theta})(\widehat{\theta} - E_{Y|\theta}\widehat{\theta})^H}_{\text{variance}} \tag{3}$$

(h) Compute $E_\theta\, b_{LMMSE}(\theta)\, b^H_{LMMSE}(\theta)$.

(i) Compute $E_\theta E_{Y|\theta}\,(\widehat{\theta}_{LMMSE} - E_{Y|\theta}\widehat{\theta}_{LMMSE})(\widehat{\theta}_{LMMSE} - E_{Y|\theta}\widehat{\theta}_{LMMSE})^H$.

Note that the sum of the positive definite matrices in (h) and (i) yields $R^{LMMSE}_{\widetilde{\theta}\widetilde{\theta}}$, for which (e) holds. Hence, in spite of the the fact that LMMSE introduces a (conditional) bias, it allows to reduce the variance so much that the sum of variance and squared bias gets lower than the variance in the unbiased case.

2. Fisher Information

Consider the data model $y_i \sim \mathcal{U}[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$ i.i.d. for $i = 1, \ldots, n$.
(Example 2 in ML estimation).

(a) Compute the scalar FIM $J(\theta)$. Compute the CRB.

(b) Interpret this unusual result. Can you relate it to what happens in ML estimation?

3. Method of Moments on a Gaussian Mixture Model

Consider i.i.d. $y_i, i = 1, \ldots, n$, with $f(y|\theta)$ a Gaussian mixture distribution with mixture parameter α, so we have

$$f(y|\theta) = (1-\alpha)\phi_1(y) + \alpha\phi_2(y)\ , \quad \phi_k(y) = \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{y^2}{2\sigma_k^2}}, k = 1, 2$$

with the parameter vector being $\theta = [\alpha\ \sigma_1^2\ \sigma_2^2]^T$.
We note that for a zero mean Gaussian variable x with variance σ^2, we get for the moments

$$\begin{array}{rcl} E\, x^{2k} & = & \frac{(2k-1)!}{2^{k-1}\,(k-1)!}\,\sigma^{2k}\ , \\ E\, x^{2k+1} & = & 0 \end{array}$$

for integer k.

(a) Compute the moments $\mu_{2k} = E_{Y|\theta} y^{2k}$.

(b) Based on this, suggest a method of moments to estimate the parameter vector θ from the $y_i, i = 1, \ldots, n$.

(c) Will the resulting estimate $\widehat{\theta}$ be consistent?

Turn the page please.

Parametric Spectrum Estimation

2. n-step Ahead Prediction of a Random DC Level plus AR(1) Noise

Consider the signal $y_k = \theta + v_k$, with θ random with zero mean, variance σ^2_θ and independent of the process v_k, which is an AR(1) noise: $v_k = a\, v_{k-1} + e_k$ where e_k is white noise.

(a) Find the correlation sequence $r_{yy}(i)$ of y_k.

(b) We shall now consider n-step ahead prediction, which means LMMSE estimation of the current sample y_k in terms of a sample y_{k-n} n sampling periods ago:

$$\widehat{y}_k \;=\; a_n\, y_{k-n}\ . \tag{4}$$

Find the optimal prediction coefficient a_n and express it in terms of σ^2_θ, σ^2_v, a and n.

(c) Find the corresponding prediction error variance $\sigma^2_{\widetilde{y}}(n)$ and express it in terms of σ^2_θ, σ^2_v, a and n.

(d) Show that $\lim\limits_{n\to\infty} \sigma^2_{\widetilde{y}}(n) = \sigma^2_v + \dfrac{\sigma^2_v \sigma^2_\theta}{\sigma^2_v + \sigma^2_\theta}$.

(e) One can observe the following bounds

$$\sigma^2_v \;\leq\; \lim_{n\to\infty} \sigma^2_{\widetilde{y}}(n) \;\leq\; \min\left(2\sigma^2_v, \sigma^2_v + \sigma^2_\theta\right)\ .$$

Explain the origin of the lower bound. And explain how the two upper bounds correspond to two specific choices of the prediction coefficient a_n.

(f) Is the signal y_k (e.g. mean) ergodic? Explain.

Turn the page please.

Wiener and Adaptive Filtering

3. FIR Equalization of a 2-Tap Channel

Consider the output of a 2-tap channel:

$$y_k \;=\; \mathrm{C}(q)\, a_k + v_k \;\;, \quad \mathrm{C}(z) \;=\; c_0 + c_{N-1}\, z^{-(N-1)} \tag{5}$$

where a_k is the transmitted symbol sequence and v_k is the additive channel noise. We assume that a_k and v_k are white processes that are mutually uncorrelated.

Steepest-Descent Algorithm

(a) Show that the correlation sequence of y_k is of the form

$$r_{yy}(n) = \begin{cases} \alpha & , n = 0 \\ \beta & , n = \pm(N-1) \\ 0 & , \text{ elsewhere} \end{cases} \tag{6}$$

for certain α and β that you will specify in terms of c_0, c_{N-1}, σ_a^2 and σ_v^2.
Hint: it is easy to find $\mathrm{S}_{yy}(z)$, which is the z transform of $r_{yy}(n)$.

(b) This same process y_k is now used as the input to an FIR filter of length $N > 2$. Give the $N \times N$ covariance matrix R_{YY} of the input signal.

(c) Find the N eigenvectors V_i and corresponding eigenvalues λ_i of R_{YY}.
Hint: $N-2$ eigenvectors are of the form $[0 \;\; * \cdots * \;\; 0]^T$ while the other two eigenvectors are of the form $[* \; 0 \cdots 0 \; *]^T$ where $*$ denotes a non-zero (in general) scalar.

(d) What is the maximum stepsize μ for convergence?

(e) What is the stepsize value μ for fastest convergence?

(f) With this fastest stepsize, what is the slowest mode?

(g) With this fastest stepsize, how fast are the other modes?

FIR Equalization

(h) Consider now FIR Equalization (Wiener filtering) with y_k as input signal and $x_k = a_k$ as desired-response signal (0 delay equalization).
Compute the N coefficients of the optimal FIR equalizer (Wiener filter) H^o, in terms of c_0, c_{N-1}, σ_a^2 and σ_v^2.

(i) Compute the associated $MMSE$.

(j) Are there values of c_0, c_{N-1} for which you would recommend using another delay d such that $x_k = a_{k-d}$? For which values of c_0, c_{N-1} and which delay d?