EURECOM 10/02/2012

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.

Bayesian Parameter Estimation

1. Bayesian Sparse Channel Estimation in Gaussian Noise with Laplacian Priors

Let the parameter vector $\theta = [\theta_1 \cdots \theta_n]^T$ contain the impulse response of an FIR channel. In wireless communications, it is common to model a fading channel impulse response as Rayleigh fading, which means that the θ_i are modeled as independent complex Gaussians. The main impact is at the level of the first and second order statistics: the θ_i are zero mean, uncorrelated and with variances $\sigma^2_{\theta_i}$. The sequence $\{\sigma^2_{\theta_i},\ i = 1, \ldots, n\}$ is called the Power Delay Profile (PDP). The PDP indicates which are the channel response coefficients that are important. However, a recent trend in estimation is *compressed sensing*, which is geared towards estimation problems with little data, in which typically Laplacian priors are used. For the sake of simplicity, we shall only consider real quantities here.

When some of the transmitted symbols are fixed (they are called "pilots"), part of the received signal may look like

$$Y \;=\; H\,\theta + V \tag{1}$$

where we shall for simplicity consider an exactly determined system: $Y = [y_1 \cdots y_n]^T$, similarly for V and H is a square $n \times n$ matrix that depends on the pilot symbols. As we know, an exactly determined system means a bad situation for estimation. In the context of an OFDM system with a judicious choice of the pilot subcarriers, it is possible to obtain an H that is an orthogonal matrix: $H^T H = I_n$ or hence $H^{-1} = H^T$. So, the independent θ_i have a Laplacian prior distribution

$$f_{\theta_i}(x) \;=\; \mu_i\, e^{-\lambda_i\,|x|} \;\;,\;\; -\infty < x < \infty\,,\; i = 1, \ldots, n \tag{2}$$

and the measurement noise is i.i.d. and Gaussian: $v_i \sim \mathcal{N}(0, \sigma_v^2)$, and independent of θ. The quantities λ_i and σ_v^2 are assumed known.

(a) Given the λ_i, determine the μ_i.

(b) Determine the prior means m_{θ_i}.

(c) Determine the prior variances $\sigma^2_{\theta_i}$ as a function of the λ_i.

(d) We can exploit the orthogonality of H to transform the problem to

$$Y' = H^T Y = H' \theta + V' . \tag{3}$$

Show that $H' = I_n$ and that $V' \sim \mathcal{N}(0, \sigma_v^2 I_n)$.

Hence equivalently $y'_i = \theta_i + v_i'$, $i = 1, \ldots, n$.

(e) Assume for questions (e) to (g) that the channel impulse response prior is Gaussian, i.e. the θ_i are independent with $\theta_i \sim \mathcal{N}(0, \sigma_{\theta_i}^2)$. Find the MAP estimator $\widehat{\theta}_i$ from the data Y'. Note that it will be of the form $\widehat{\theta}_i = \alpha_i \, y_i'$ for some α_i that you need to specify.

(f) Is $\widehat{\theta}_i$ in (e) unbiased in the Bayesian sense? Is it conditionally unbiased (in the deterministic sense)?

(g) Compute the estimation error MSE $\sigma^2_{\widetilde{\theta}_i}$ corresponding to (e).

(h) For the remaining questions, assume again the Laplacian prior for the channel response θ. Compute the MAP estimator of θ_i given Y'.
Hints: express $y_i' = |y_i'| \operatorname{sign}(y_i')$, $\theta_i = |\theta_i| \operatorname{sign}(\theta_i)$. First find $\operatorname{sign}(\widehat{\theta}_i)$ and then $|\widehat{\theta}_i|$, for which two cases can occur.

Note that the Laplacian prior leads for a portion of the channel coefficient estimates to $\widehat{\theta}_i = 0$, the portion of which increases as SNR lowers. This means that equivalently a reduced number of the channel coefficients only get actually estimated.

(i) Is $\widehat{\theta}_i$ in (h) unbiased (in a Bayesian sense)?

(j) Compute the CRB for $\widehat{\theta}_i$.
Note that $\ln f(y_i', \theta_i) = \ln f(y_i'|\theta_i) + \ln f(\theta_i)$ and that $\int \delta(x) \, f(x) \, dx = f(0)$.

Turn the page please.

Parametric Spectrum Estimation

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

Consider the same signal $y_k = \theta + v_k$ as in the previous problem, with v_k AR(1) noise, but let θ now be random with zero mean, variance σ_θ^2 and independent of the process v_k.

(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_θ^2, σ_v^2, a and n.

(c) Find the corresponding prediction error variance $\sigma_{\tilde{y}}^2(n)$ and express it in terms of σ_θ^2, σ_v^2, a and n.

(d) Show that $\lim_{n\to\infty} \sigma_{\tilde{y}}^2(n) = \sigma_v^2 + \dfrac{\sigma_v^2 \sigma_\theta^2}{\sigma_v^2 + \sigma_\theta^2}$.

(e) One can observe the following bounds

$$\sigma_v^2 \; \leq \; \lim_{n\to\infty} \sigma_{\tilde{y}}^2(n) \; \leq \; \min\left(2\sigma_v^2, \sigma_v^2 + \sigma_\theta^2\right) \; .$$

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:

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

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?