EURECOM 24/06/2011
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 $\theta_i$ are modeled as independent complex Gaussians. The main impact is at the level of the first and second order statistics: the $\theta_i$ 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 $\theta_i$ 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 $\theta$. The quantities $\lambda_i$ and $\sigma_v^2$ are assumed known.

(a) Given the $\lambda_i$, determine the $\mu_i$.

(b) Determine the prior means $m_{\theta_i}$.

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

(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 $\theta_i$ are independent with $\theta_i \sim \mathcal{N}(0, \sigma^2_{\theta_i})$. 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 $\alpha_i$ 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 $\theta$. Compute the MAP estimator of $\theta_i$ given $Y^{'}$.
Hints: express $y_i^{'} = |y_i^{'}|\,\mathrm{sign}(y_i^{'})$, $\theta_i = |\theta_i|\,\mathrm{sign}(\theta_i)$. First find $\mathrm{sign}(\widehat{\theta}_i)$ 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.

## Spectrum Estimation

**2. Linear Prediction of the Sum of an Oscillating Signal and a DC Offset**

Consider the following random signal

$$y_k = s_k + v_k \quad . \tag{4}$$

The signal $s_k$ is constant but random. This means that $s_k \equiv A \;\; \forall k$, where $A$ is a random variable with zero mean and variance $\sigma_s^2$. The signal $v_k$ is oscillating and random. This means that $v_k = C\,(-1)^k$ , $\forall k$, where $C$ is a random variable, independent of $A$, with zero mean and variance $\sigma_v^2$.

**(i)** Give the autocorrelation function, $r_i = \mathrm{E}\;\, y_{k+i} y_k$, in terms of $\sigma_s^2$ and $\sigma_v^2$.

**(ii)** Express the autocorrelation matrix, $R_n$, as the sum of a constant matrix and an oscillating Toeplitz matrix. Show that both matrices are of rank one (i.e. of the form $S\,S^T$ where $S$ is a column vector). Express the two rank one matrices in terms of the vectors $\mathbf{1}_N = [1 \;\; 1 \cdots 1]^T$ and $\overline{\mathbf{1}}_n = [1 \;\; (-1) \cdots (-1)^{n-1}]^T$ respectively.

**(iii)** Consider linear prediction of $y_k$ of order 1. Compute the prediction coefficient $A_{1,1}$ and the prediction error variance $\sigma_{f,1}^2$ as a function of $\sigma_s^2$ and $\sigma_v^2$.

**(iv)** Consider linear prediction of $y_k$ of order 2. Compute the prediction coefficients $A_{2,i}$, $i = 1, 2$ and the prediction error variance $\sigma_{f,2}^2$ as a function of $\sigma_s^2$ and $\sigma_v^2$. What do you conclude?

**(v)** Give an expression for the predicted value $\widehat{y}_k = y_k - f_{2,k}$ and interpret the result obtained.

**(vi)** Give the prediction coefficients $A_{n,i}$ for $n > 2$.

**(vii)** Find the PARCORs $K_n$ as a function of $n$ (for $n = 1, 2, ...$).

Turn the page please.

## Wiener Filtering and Equalization

**3. FIR (U)MMSE Linear Equalization of a FIR Channel**

Consider a causal FIR equalizer $H$ with $N$ coefficients, $\widehat{x}_k = H^T\,Y_k$. For a FIR channel of length $L$, the received signal vector $Y_k$ can be written as

$$\begin{bmatrix} y_k \\ y_{k-1} \\ \vdots \\ y_{k-N+1} \end{bmatrix} = \begin{bmatrix} c_0 & c_1 & \cdots & c_{L-1} & 0 & \cdots & 0 \\ 0 & c_0 & \cdots & c_{L-2} & c_{L-1} & \ddots & 0 \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & c_0 & \cdots & \cdots & c_{L-1} \end{bmatrix} \begin{bmatrix} s_k \\ s_{k-1} \\ \vdots \\ s_{k-N-L+2} \end{bmatrix} + \begin{bmatrix} v_k \\ v_{k-1} \\ \vdots \\ v_{k-N+1} \end{bmatrix}$$

or $\underbrace{Y_k}_{N\times 1} = \underbrace{C}_{N\times(N+L-1)}\;\underbrace{S_k}_{(N+L-1)\times 1} + \underbrace{V_k}_{N\times 1} = [\underline{c}_0\;\underline{c}_1\;\cdots\;\underline{c}_{N+L-2}]\;S_k + V_k = \sum_{i=0}^{N+L-2}\underbrace{\underline{c}_i}_{N\times 1}\;s_{k-i} + V_k$ . The vector $\underline{c}_i$ is column $i{+}1$ of the matrix $C$, so $\underline{c}_i = C\,\underline{e}_i$ where $\underline{e}_i$ is a standard unit vector with all zeros and a 1 in position $i{+}1$. The symbol sequence $s_k$ is considered to be white noise with zero mean and variance $\sigma_s^2$. The additive noise $v_k$ is independent of the symbol sequence and white Gaussian with zero mean and variance $\sigma_v^2$. We shall see that it may be advantageous to introduce an equalization delay $d$. Hence consider $x_k^{(d)} = s_{k-d},\; d \in \{0, 1, \ldots, N{+}L{-}2\}$. MMSE FIR equalization is a particular instance of FIR Wiener filtering. Hence the MMSE FIR equalizer coefficients $H_{MMSE}^{(d)}$ satisfy the normal equations

$R_{YY}\,H_{MMSE}^{(d)} = R_{Yx^{(d)}}$ or hence $H_{MMSE}^{(d)} = R_{YY}^{-1}\,R_{Yx^{(d)}}$, and

the MMSE is $\sigma^2_{\widetilde{x}^{(d)}_{MMSE}} = R_{x^{(d)}x^{(d)}} - R_{x^{(d)}Y}R_{YY}^{-1}\,R_{Yx^{(d)}}$.

(a) Determine $R_{YY}$ in terms of $\sigma_s^2$, $\sigma_v^2$, the matrix $C$ and the identity matrix $I_N$, and determine $R_{Yx^{(d)}}$ in terms of $\sigma_s^2$ and the vector $\underline{c}_d$.

(b) Express the MMSE $\sigma^2_{\widetilde{x}^{(d)}_{MMSE}}$ in terms of these same quantities.

The corresponding (naive) SNR is $\mathrm{SNR}_{MMSE}^{(d)} = \dfrac{\sigma_s^2}{\sigma^2_{\widetilde{x}^{(d)}_{MMSE}}}$.

In what follows, we shall consider the specific case of a channel with $L = 2$ coefficients $c_0 = 1$, $c_1 = -a$, and no noise: $\sigma_v^2 = 0$. The range of possible delays is now limited to $d \in \{0, 1, \ldots, N\}$. In the absence of noise, the MSE is determined by intersymbol interference which is unavoidable here with an FIR equalizer, so the MSE is nonzero even in absence of noise.

(c) Show that MMSE $\sigma^2_{\widetilde{x}^{(d)}_{MMSE}} = \sigma_s^2\;\underline{e}_d^T P^{\perp}_{C^T}\underline{e}_d$ where the projection matrices are $P_X^{\perp} = I - P_X$ and $P_X = X(X^TX)^{-1}X^T$.

(d) Consider the channel transfer function $\mathrm{C}(z) = 1 - a\,z^{-1}$. Hence $\mathrm{C}(z = a) = 0$. As a result, show that

$$C\,V = 0 \tag{5}$$

where $V = [1\;\frac{1}{a}\;\frac{1}{a^2}\cdots\frac{1}{a^N}]^T$. So, since $C^T$ has full column rank, we get $P^{\perp}_{C^T} = P_V$.

(e) Now compute again MMSE $\sigma^2_{\widetilde{x}^{(d)}_{MMSE}}$ for $d \in \{0, 1, \ldots, N\}$, as a function of $a$, $N$ and $d$.

(f) Which is the optimal delay $d$ as a function of $a$?