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

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

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 an equivalent form of (3) is: $y'_i = \theta_i + v_i^{'}\,,\; i = 1, \ldots, n$.

(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 $y_i^{'} = |y_i^{'}|\; \mathrm{sign}(y_i^{'})$, $\theta_i = |\theta_i|\, \mathrm{sign}(\theta_i)$. 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 only a reduced number of the channel coefficients 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.

Deterministic Parameter Estimation

2. Maximum Likelihood & Fisher Information

Consider the data model $y_i \sim \mathcal{N}(0, \sigma^2)$ i.i.d. for $i = 1, \ldots, n$ with unknown variance $\theta = \sigma^2$.
(See BLUE example in Lecture 5).

(a) Compute the Maximum Likelihood estimator $\widehat{\sigma^2}_{ML}$.

(b) How does the ML estimator $\widehat{\sigma^2}_{ML}$ compare to the BLUE estimator $\widehat{\sigma^2}_{BLUE}$?

(c) Is this ML estimator unbiased? (derive the bias)

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

(e) Is this ML estimator consistent? Why?

(f) Is the ML estimator a UMVUE in this case? Why?

Spectrum Estimation

3. Spectral Resolution Issues

We consider the use of the Blackman-Tukey (BT) spectral estimator and more specifically with a Bartlett window of length $2M+1$. We want a spectral resolution in normalized frequency of at least 0.02 (assume the common rule of thumb for the resolution of two equiamplitude sinusoids). We furthermore want the BT spectral estimator to have a variance that is at least nine times lower than the variance of the periodogram. What is the minimum number N of samples that we need to have to reach these specifications?

4. Time-Frequency Signal Analysis Resolution Issues

(a) Assume we use the discrete-time Short-Time Fourier Transform (DFT) to do time-frequency analysis. Given a certain frequency resolution (Δf) and a certain temporal resolution (Δt), is it possible to find a unique sampling frequency f_s and number of frequency bins N that lead to these given frequency and temporal resolutions?

(b) Doing again time-frequency analysis using the DFT and given a sampling frequency $f_s = 16$kHz and a temporal resolution $\Delta t = 1$ms, what is the frequency resolution Δf and what is the number of frequency bins N?

(c) If we now use the discrete-time Wavelet Transform (DTWT) and we use the same number of subbands as in (b), what are the frequency resolution Δf and the temporal resolution Δt in

  (i) the highest subband?
  (ii) the lowest subband?

Turn the page please.

Wiener Filtering

5. Noise Canceling via Wiener Filtering

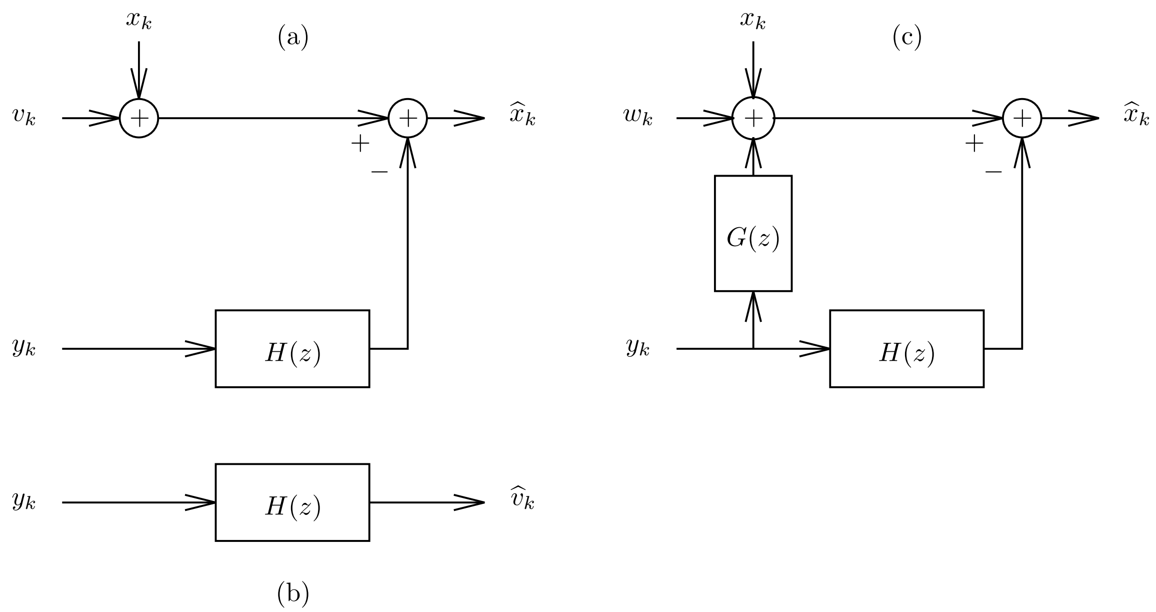

Figure 1: Noise canceling problem.

In Fig. 1(a), the noise canceling paradigm has been sketched. All processes involved are zero mean and stationary. The signal of interest, x_k, can only be measured with additive noise v_k, leading to the measured signal x_k+v_k. On the other hand, we can also measure another signal y_k that is correlated with the measurement noise v_k, but uncorrelated with x_k ($S_{xy}(z) = 0$). We shall filter this signal y_k with the filter $H(z) = \sum_k h_k\, z^{-k}$, and subtract the filter output from the measured signal x_k+v_k to form a less noisy estimate of x_k, namely $\widehat{x}_k$.

(i) Show that the filter $H(z)$ that minimizes the MSE $\mathrm{E}\,(\widehat{x}_k - x_k)^2$ is the same as the filter $H(z)$ that follows from the Wiener filtering problem shown in Fig. 1(b) in which $H(z)$ minimizes the MSE $\mathrm{E}\,(\widehat{v}_k - v_k)^2$ where $\widehat{v}_k = h_k * y_k$ ($*$ denotes convolution).

(ii) Find $H(z)$ that solves the problems mentioned in **(i)**.

(iii) Find the associated MSE $\mathrm{E}\,(\widehat{x}_k - x_k)^2$.

(iv) A particular case is depicted in Fig. 1(c) where $v_k = g_k * y_k + w_k$. $G(z) = \sum_k g_k\, z^{-k}$ is some fixed filter and w_k is uncorrelated with x_k and y_k and has variance σ_w^2. What does the optimal filter $H(z)$ become in this case (that leads to the MMSE estimate $\widehat{x}_k$)?

(v) Using this optimal $H(z)$, what is $\widehat{x}_k$ equal to?
What is the associated MSE $\mathrm{E}\,(\widehat{x}_k - x_k)^2$?