Exercise sheet 4

Exercise 4 - 1

Consider the potential $V(\vec{r})$ which is symmetric with respect to the radial distance $r = |\vec{r}|$,

$$V(r) = a\sqrt{r} + b. \tag{1}$$

This potential is parametrized by the unknown numbers $a, b \in \mathbb{R}$ and can be measured at strictly positive radii, i.e., $r > 0$. Furthermore, only a single data point $d \in \mathbb{R}$ can be obtained,

$$d = V(r) + n, \tag{2}$$

where the noise n is assumed to obey a Gaussian statistic $\mathcal{P}(n) = \mathcal{G}(n, N)$. The noise variance $N = N(r)$, however, depends on the measurement position,

$$N(r) = r^2 + 3. \tag{3}$$

a) Find an expression for the information entropy $S\left[\mathcal{P}(s|d)\right]$ for a Gaussian posterior

$$\mathcal{P}(s|d) = \mathcal{G}(s - m, D)\ , \tag{4}$$

with mean m and covariance D (2 points).

b) Consider the signal $s = \begin{pmatrix} a \\ b \end{pmatrix}$, for which a Gaussian prior $\mathcal{P}(s) = \mathcal{G}(s, \mathbb{1})$ can be assumed. (2 points)

- Write Eq. (2) in the form $d = R\,s + n$ and give R explicitly.
- Work out an expression for the joint probability $\mathcal{P}(d, s)$ and calculate the corresponding Hamiltonian $H(d, s) = -\log \mathcal{P}(d, s) = \frac{1}{2} s^\dagger D^{-1} s - j^\dagger s + H_0$. You may drop H_0.
- Identify the information source j and the inverse information propagator D^{-1}.

c) You verified in **a)** that information entropy $S = S(D)$ is a monotonically increasing function of $|D|$. Find the best position $\tilde{r}$ to estimate both, a and b, by minimizing $|D|$ from **b)** with respect to r (1 point).

d) Now, consider the signal $s = a$ for which b becomes a nuisance parameter. (1 point)

- Work out an expression for the joint probability $\mathcal{P}(d, a)$, and calculate the corresponding Hamiltonian $H(d, a) = -\log \mathcal{P}(d, a) = \frac{1}{2} D^{-1} a^2 - ja + H_0$. You may drop H_0.
- Identify the information source j and the information propagator D.

e) Find the best position $\tilde{r}_a$ to estimate a irrespectively of b, by minimizing D from **d)** with respect to r (1 point).

f) Guess at which radius $\tilde{r}_b$ one should measure in order to obtain the most certain estimate for the parameter b. No justification required (1 point).

Exercise 4 - 2

It was shown in the lecture that for arbitrary signal-, noise-, and data-statistics with known correlations $\left\langle ss^\dagger \right\rangle_{(s,d)}$, $\left\langle ds^\dagger \right\rangle_{(s,d)}$, and $\left\langle dd^\dagger \right\rangle_{(s,d)}$, the optimal linear filter is given by

$$m = \left\langle sd^\dagger \right\rangle_{(s,d)} \left\langle dd^\dagger \right\rangle_{(s,d)}^{-1} d. \tag{5}$$

A linear response matrix R and a noise covariance matrix N can be *defined* via the following identifications:

$$\left\langle ss^\dagger \right\rangle_{(s,d)} \equiv S \tag{6}$$
$$\left\langle ds^\dagger \right\rangle_{(s,d)} \equiv RS \tag{7}$$
$$\left\langle dd^\dagger \right\rangle_{(s,d)} \equiv RSR^\dagger + N \tag{8}$$

Find expressions for R and N in terms of $\left\langle ss^\dagger \right\rangle_{(s,d)}$, $\left\langle ds^\dagger \right\rangle_{(s,d)}$, and $\left\langle dd^\dagger \right\rangle_{(s,d)}$(2 points).

Exercise 4 - 3

You are interested in three numbers, $s = (s_1, s_2, s_3) \in \mathbb{R}^3$. Your measurement device, however, only measures three differences between the numbers, according to

$$d_1 = s_1 - s_2 + n_1 \tag{9}$$
$$d_2 = s_2 - s_3 + n_2 \tag{10}$$
$$d_3 = s_3 - s_1 + n_3 \tag{11}$$

with some noise vector $n \in \mathbb{R}^3$. Assume a Gaussian prior $\mathcal{P}(s) = \mathcal{G}(s, S)$ for s and a Gaussian PDF for the noise, $\mathcal{P}(n) = \mathcal{G}(n, N)$, with $N_{ij} = \sigma^2 \delta_{ij}$.

a) Assume that the prior is degenerate, i.e., $S^{-1} \equiv 0$. Write down the response matrix, try to give the posterior $\mathcal{P}(s|d)$, and explain why this is problematic (2 points).

b) Now assume that $S_{ij} = \sigma^2 \delta_{ij}$. Work out the posterior $\mathcal{P}(s|d)$ in this case (1 point).

Note: Using a computer algebra system, e.g., SAGE (`http://www.sagemath.org/`), for the matrix operations is okay.

Exercise 4 - 4

You have conducted a measurement of a quantity at n positions $\{x_i\}_i$, yielding n data points $\{(x_i, d_i)\}_i$. Now you want to fit some function to these data points. To this end, you write the function as a linear combination of m basis functions $\{f_j(x)\}_j$, i.e.,

$$f(x) = \sum_{j=1}^{m} s_j f_j(x). \tag{12}$$

If, for example, you were to fit a second order polynomial, you could choose the monomials as basis functions, i.e., $f(x) = s_2 x^2 + s_1 x + s_0$.
The fitting process now comes down to determining the coefficients $\{s_j\}_j$, allowing for some Gaussian and independent measurement error, i.e.,

$$d_i = \sum_{j=1}^{m} s_j f_j(x_i) + n_i. \tag{13}$$

Assume that you do not know anything about the coefficients a priori, i.e., $S^{-1} \equiv 0$, where $S_{ik} = \langle s_i s_k \rangle_{\mathcal{P}(s)}$.

a) Write down the response matrix for this problem (1 point).

b) For a given set of m basis functions, how many data points n are at least necessary for the calculation of the posterior mean of the coefficients (2 points)?

c) Now let's make a linear fit. Assuming $N_{ik} = \langle n_i n_k \rangle_{\mathcal{P}(n)} = \eta^{-1}\delta_{ik}$, choose two basis functions and work out the explicit formula for the posterior mean of the two coefficients (3 points).

Exercise 4 - 5

Assume that a quantity y is linearly dependent on a quantity x, i.e., $y(x) = a + bx$. Assume further that the quantity y has been measured at $m-1$ different positions $(x_i)_i,\ i = 1, \ldots, m-1$, subject to additive uncorrelated Gaussian noise, i.e.,

$$d_i = y(x_i) + n_i, \quad n \hookleftarrow \mathcal{G}(n, N), \quad N_{ij} = \delta_{ij}\sigma_i^2. \tag{14}$$

Assuming a Gaussian prior for the parameters a and b, i.e.,

$$s = \begin{pmatrix} a \\ b \end{pmatrix} \hookleftarrow \mathcal{G}(s, S), \quad S = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, \tag{15}$$

a linear fit can be performed using Wiener filter theory.

You have enough money left to finance one additional measurement with uncertainty σ_m. How should you choose the position x_m for that measurement to gain optimal knowledge about the parameter a?

Hint: Use the quadratic loss function $\mathcal{L}(s, x_m) = (a - m_a)^2$ and the formalism of risk minimization. Here, m_a is the Wiener filter estimate after the m measurements (4 points).