Solutions 8

Solution 1. Firing Neuron

(a) Random variables $N_i,\ i = 1, \dots, 20$ obey the Poisson distribution given by

$$P(N_i = k) = \frac{\lambda^k e^{-\lambda}}{k!}\,.$$

Note that the parameter t has been omitted since each interval i is one second long.

Since the activity of the recorded neuron is modeled by a Poison process, and random variables $N_i,\ i = 1, \dots, 20$ model the number of spikes in non-overlapping intervals, they are independent. Therefore

$$\begin{aligned}
&P(N_1 = n_1, \dots, N_{20} = n_{20}) \\
&= P(N_1 = n_1) \cdot \ldots \cdot P(N_{20} = n_{20}) \\
&= \frac{\lambda^{n_1} e^{-\lambda}}{n_1!} \cdot \ldots \cdot \frac{\lambda^{n_{20}} e^{-\lambda}}{n_{20}!}\,.
\end{aligned}$$

Random variable N obeys the Poisson distribution as well:

$$P(N = n) = \frac{(20\lambda)^n e^{-20\lambda}}{n!}\,.$$

(b) Denote by n the number of recorded spikes in the entire experiment (in our case, $n = 100$). The Poisson process likelihood function is a function of the parameter λ (firing rate) given by the probability that random variable N takes the value n:

$$f_N(n) = f_N(N = n;\, \lambda) = \frac{(20\lambda)^n e^{-20\lambda}}{n!}\,.$$

(c) Maximizing the likelihood function $f_N(n;\, \lambda)$ is equivalent to maximizing its log-likelihood $\ln f_N(n;\, \lambda)$. The value of the parameter λ is obtained by taking the derivative of the log-likelihood with respect to λ and equating it to zero:

$$\frac{\partial \ln f_N(n;\, \lambda)}{\partial \lambda} = \frac{n}{\lambda} - 20 = 0\,.$$

It follows that $\lambda = \frac{n}{20}$.

(d) In order to be able to characterize the Markov chain, we need to estimate two state probabilities: π_3 and π_8, and four transition probabilities: p_{33}, p_{38}, p_{83}, and p_{88}. However, since state probabilities and transition probabilities from each state should sum to one, we know that three parameters: π_8, p_{38} and p_{83} are dependent, and need not be determined.

(e) The Markov chain likelihood function is a function of a set Θ of Markov chain parameters (state and transition probabilities), and is given by the probability that a sequence of Markov chain's states is equal to the given sequence:

$$f_{\Lambda[1],\ldots,\Lambda[20]}(\lambda_1,\ldots,\lambda_{20};\,\Theta) = P(\Lambda[1]=\lambda_1,\ldots,\Lambda[20]=\lambda_{20};\,\Theta)$$

Using the Markov chain property, we obtain

$$\begin{aligned} f_{\Lambda[1],\ldots,\Lambda[20]}(\lambda_1,\ldots,\lambda_{20};\,\Theta) &= P(\Lambda[20]=\lambda_{20}|\Lambda[19]=\lambda_{19}) \times \ldots \\ &\times P(\Lambda[2]=\lambda_2|\Lambda[1]=\lambda_1)\times \\ &\times P(\Lambda[1]=\lambda_1) \\ &= p_{\lambda_{20}\lambda_{19}} \cdot \ldots \cdot p_{\lambda_2\lambda_1}\pi_{\lambda_1}\,. \end{aligned}$$

(f) We start by maximizing the log-likelihood $\ln f_{\Lambda[1],\ldots,\Lambda[20]}(\lambda_1,\ldots,\lambda_{20};\,\Theta)$ subject to state and transition probabilities' constraints:

$$\begin{aligned} J &= \ln p_{\lambda_{20}\lambda_{19}} + \ldots + \ln p_{\lambda_2\lambda_1} + \ln \pi_{\lambda_1} \\ &\quad - \alpha_1(\pi_3+\pi_8-1) - \alpha_2(p_{33}+p_{38}-1) - \alpha_3(p_{83}+p_{88}-1)\,. \end{aligned}$$

Using the observed data, we have

$$\begin{aligned} J &= 5\ln p_{33} + 6\ln p_{38} + 6\ln p_{83} + 2\ln p_{88} + \ln \pi_3 \\ &\quad - \alpha_1(\pi_3+\pi_8-1) - \alpha_2(p_{33}+p_{38}-1) - \alpha_3(p_{83}+p_{88}-1)\,. \end{aligned}$$

Taking partial derivatives of J w.r.t. state and transition probabilities and equating them to zero, one obtains the following estimation for Markov chain parameters:

$$\hat{\pi}_3 = 1,\ \hat{\pi}_8 = 0,\ \hat{p}_{33} = \frac{5}{11},\ \hat{p}_{38} = \frac{6}{11},\ \hat{p}_{83} = \frac{3}{4},\ \hat{p}_{88} = \frac{1}{4}\,.$$

Solution 2.

(a) We have:

$$\begin{aligned} X_0[n] &= V_0[n] + V_0[n-1] + S[n] = \\ &= \begin{cases} V_0[n] + V_0[n-1] + V_1[n], & \text{n is even} \\ V_0[n] + V_0[n-1] + V_2[n], & \text{n is odd} \end{cases} \end{aligned}$$

Process $X_0[n]$ is a sum of two processes $D[n]$ and $S[n]$ that are Gaussian at every instant. Therefore, $X_0[n]$ is Gaussian, as well. To check if the process is wide sense stationary we compute the mean and the variance.

$$\mathbb{E}[X_0[n]] = 0$$

$$\begin{aligned} \mathbb{E}[X_0[n]X_0[n]] &= \mathbb{E}[(V_0[n] + V_0[n-1] + S[n])(V_0[n] + V_0[n-1] + S[n])] \\ &= \begin{cases} 2\sigma_{V_0}^2 + \sigma_{V_1}^2, & \text{n is even} \\ 2\sigma_{V_0}^2 + \sigma_{V_2}^2, & \text{n is odd} \end{cases} \end{aligned}$$

We can see that $S[n]$ is not a wss process and consequently $X_0[n]$ is not wss.

In order to compute the correlation of $X_0[n]$ we need first to compute the correlation of $S[n]$.

$$R_S[n,m] = \mathbb{E}[S[n]S[m]] = \begin{cases} \mathbb{E}[V_1[n]V_1[m]] = \delta[n-m]\sigma_{V_1}^2, & \text{n and m are even} \\ \mathbb{E}[V_1[n]V_2[m]] = 0, & \text{n even, m odd} \\ \mathbb{E}[V_2[n]V_1[m]] = 0, & \text{n odd, m even} \\ \mathbb{E}[V_2[n]V_2[m]] = \delta[n-m]\sigma_{V_2}^2, & \text{n and m are odd} \end{cases}$$

Then,

$$\begin{aligned} R_{X_0}[n,m] &= \mathbb{E}[X_0[n]X_0[m]] \\ &= \mathbb{E}[(V_0[n]+V_0[n-1]+S[n])(V_0[m]+V_0[m-1]+S[m])] \\ &= 2\sigma_{V_0}^2\delta[n-m] + \sigma_{V_0}^2\delta[n-1-m] + \sigma_{V_0}^2\delta[n-m+1] + R_S[n,m] \\ &= \begin{cases} 2\sigma_{V_0}^2\delta[n-m] + \sigma_{V_1}^2\delta[n-m], & \text{n and m are even} \\ 2\sigma_{V_0}^2\delta[n-m] + \sigma_{V_2}^2\delta[n-m], & \text{n and m are odd} \\ \sigma_{V_0}^2\delta[n-1-m] + \sigma_{V_0}^2\delta[n-m+1], & \text{otherwise} \end{cases} \end{aligned}$$

(b) We define the cost function as:

$$J_{min} = \mathbb{E}[|D[n] - \sum_{k=0}^{1} f_n(k)X_0[n-k]|^2].$$

The optimal filter is given by:

$$\begin{bmatrix} f_n(0) \\ f_n(1) \end{bmatrix} = \begin{bmatrix} R_{X_0}[n,n] & R_{X_0}[n-1,n] \\ R_{X_0}[n,n-1] & R_{X_0}[n-1,n-1] \end{bmatrix}^{-1} \begin{bmatrix} R_{DX_0}[n,n] \\ R_{DX_0}[n,n-1] \end{bmatrix}$$

where

$$R_{DX_0}[n,m] = \mathbb{E}[D[n]X_0[m]] = \mathbb{E}[D[n](D[m]+S[m])] = R_D[n,m] + R_{DS}[n,m],$$

$$R_D[n,m] = 2\sigma_{V_0}^2\delta[n-m] + \sigma_{V_0}^2\delta[n-1-m] + \sigma_{V_0}^2\delta[n-m+1],$$

$$R_{DS}[n,m] = 0 \quad \text{for all n and m}$$

Then, when n is even we have:

$$\begin{bmatrix} f_n(0) \\ f_n(1) \end{bmatrix} = \begin{bmatrix} 2\sigma_{V_0}^2 + \sigma_{V_1}^2 & \sigma_{V_0}^2 \\ \sigma_{V_0}^2 & 2\sigma_{V_0}^2 + \sigma_{V_2}^2 \end{bmatrix}^{-1} \begin{bmatrix} 2\sigma_{V_0}^2 \\ \sigma_{V_0}^2 \end{bmatrix} = \begin{bmatrix} 7/11 \\ 1/11 \end{bmatrix},$$

and when n is odd we have:

$$\begin{bmatrix} f_n(0) \\ f_n(1) \end{bmatrix} = \begin{bmatrix} 2\sigma_{V_0}^2 + \sigma_{V_2}^2 & \sigma_{V_0}^2 \\ \sigma_{V_0}^2 & 2\sigma_{V_0}^2 + \sigma_{V_1}^2 \end{bmatrix}^{-1} \begin{bmatrix} 2\sigma_{V_0}^2 \\ \sigma_{V_0}^2 \end{bmatrix} = \begin{bmatrix} 5/11 \\ 2/11 \end{bmatrix}.$$

(c) We define the cost function as:

$$J_{min} = \mathbb{E}[|S[n] - \sum_{k=0}^{1} f_n(k)X_1[n-k]|^2].$$

The optimal filter is given by:

$$\begin{bmatrix} f_n(0) \\ f_n(1) \end{bmatrix} = \begin{bmatrix} R_{X_1}[n,n] & R_{X_1}[n-1,n] \\ R_{X_1}[n,n-1] & R_{X_1}[n-1,n-1] \end{bmatrix}^{-1} \begin{bmatrix} R_{SX_1}[n,n] \\ R_{SX_0}[n,n-1] \end{bmatrix}$$

where

$$\begin{aligned} R_{X_1}[n,m] &= \mathbb{E}[X_1[n]X_1[m]] \\ &= \mathbb{E}[(V_3[n]-V_3[n-1]+S[n])(V_3[m]-V_3[m-1]+S[m])] \\ &= 2\sigma_{V_3}^2\delta[n-m]-\sigma_{V_3}^2\delta[n-1-m]-\sigma_{V_3}^2\delta[n-m+1]+R_S[n,m] \\ &= \begin{cases} 2\sigma_{V_3}^2\delta[n-m]+\sigma_{V_1}^2\delta[n-m], & \text{n and m are even} \\ 2\sigma_{V_3}^2\delta[n-m]+\sigma_{V_2}^2\delta[n-m], & \text{n and m are odd} \\ -\sigma_{V_3}^2\delta[n-1-m]-\sigma_{V_3}^2\delta[n-m+1], & \text{otherwise} \end{cases}, \end{aligned}$$

$$\begin{aligned} R_{SX_1}[n,m] = \mathbb{E}[S[n]X_1[m]] &= \mathbb{E}[S[n](V_3[m]-V_3[m-1]+S[m])] \\ = R_S[n,m] &= \begin{cases} \sigma_{V_1}^2\delta[n-m] & \text{n,m even} \\ \sigma_{V_2}^2\delta[n-m] & \text{n,m odd} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

For even n we have:

$$\begin{bmatrix} f_n(0) \\ f_n(1) \end{bmatrix} = \begin{bmatrix} 2\sigma_{V_3}^2+\sigma_{V_1}^2 & -\sigma_{V_3}^2 \\ -\sigma_{V_3}^2 & 2\sigma_{V_3}^2+\sigma_{V_2}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sigma_{V_1}^2 \\ 0 \end{bmatrix} = \begin{bmatrix} 6/17 \\ 1/17 \end{bmatrix},$$

and when n is odd we have:

$$\begin{bmatrix} f_n(0) \\ f_n(1) \end{bmatrix} = \begin{bmatrix} 2\sigma_{V_3}^2+\sigma_{V_2}^2 & -\sigma_{V_3}^2 \\ -\sigma_{V_3}^2 & 2\sigma_{V_3}^2+\sigma_{V_1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sigma_{V_2}^2 \\ 0 \end{bmatrix} = \begin{bmatrix} 12/17 \\ 2/17 \end{bmatrix}.$$

The estimated of $D[n]$ can be determine as:

$$\hat{D}[n] = X_0[n] - \hat{S}[n]$$

(d) The optimal filter is obtained in the same way as before

$$\begin{bmatrix} \bar{f}_{n,0} \\ \bar{f}_{n,1} \end{bmatrix} = \begin{bmatrix} R_{\bar{X}}[n,n] & R_{\bar{X}}[n-1,n] \\ R_{\bar{X}}[n,n-1] & R_{\bar{X}}[n-1,n-1] \end{bmatrix}^{-1} \begin{bmatrix} R_{D\bar{X}}[n,n] \\ R_{D\bar{X}}[n,n-1] \end{bmatrix}$$

where

$$R_{\bar{X}}[n,m] = R\left[\begin{bmatrix} X_0[n] \\ X_1[n] \end{bmatrix} [\ X_0^*[m] \quad X_1^*[m]\ ]\right] = \begin{bmatrix} R_{X_0}[n,m] & R_{X_0X_1}[n,m] \\ R_{X_1X_0}[n,m] & R_{X_1}[n,m] \end{bmatrix},$$

and

$$R_{D\bar{X}}[n,m] = \begin{bmatrix} R_{DX_0}[n,m] \\ R_{DX_1}[n,m] \end{bmatrix}.$$