Week 06 lecture 26

- Recap of lecture 25, about cross correlation and STA.

Consider the system,

- Output of a point process being teh spike trains (P(t)), $\Delta(bin\ width)$, $M\ bins$.
- The total time duration is given by T.
- The $\Delta(bin\ width)$ is so small that there can be max of only one spike per Δ.
- The bins are 0/1.
- Then the probability of $q_i, P(q_i = 1) = \lambda(\Delta_i).\Delta$
- Then the cross correlation of pulse train to the input X,

$$R_{PX} \simeq R_{qX}$$

Since the pulse trains are not going to be for infinite duration , the cross correlation function can be an estimate ,

Let the number of spikes be N_T

$$\hat{R}_{PX} = \frac{1}{N}\int_0^T X(t - \tau)\sum_{i=1}^{N_T} \delta(t - t_i)dt$$

$= \frac{1}{N}\sum_{i=1}^{N_T} \int_0^T X(t - \tau)\delta(t - t_i)dt$

$= \frac{1}{N}\sum_{i=1}^{N_T} X(t_i - \tau)$

Consider the spike train has M bins,

- Relationship between $\lambda\ and\ x$
- From STA, we have $R_{PX}(\tau) \simeq R_{qX}(\tau)$
- $\hat{R}_{qX}(\tau)=\frac{1}{M}[E[q(i\Delta)\,X(i\Delta - \tau) \mid X(i\Delta - \tau)]$

- Impulse response is a scaled version of spike triggered average

Price's Theorem

Consider a bivariate normal distribution in variables $\mathbf{x}$ and $\mathbf{y}$ with covariance

$$\rho = \rho_{11} = \langle \mathbf{x}\, \mathbf{y} \rangle - \langle \mathbf{x} \rangle \langle \mathbf{y} \rangle$$

and an arbitrary function $g(x, y)$. Then the expected value of the random variable $g(\mathbf{x}, \mathbf{y})$

$$\langle g(\mathbf{x}, \mathbf{y}) \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y)\, d\, x\, d\, y$$

satisfies

$$\frac{\partial^n \langle g(\mathbf{x}, \mathbf{y}) \rangle}{\partial \rho^n} = \left\langle \frac{\partial^{2n} g(\mathbf{x}, \mathbf{y})}{\partial \mathbf{x}^n\, \partial \mathbf{y}^n} \right\rangle.$$

REf: Papoulis, A. "Price's Theorem and Join Moments." *Probability, Random Variables, and Stochastic Processes, 2nd ed.* New York: McGraw-Hill, pp. 226-228, 1984.

-From the linear model:- X(t) can be considered as a gaussian white noise with mean =0;

- The covariance then becomes, R_{YX}

- IN the LTI system ,

can be represented as $\lambda(y)$.

- The non linearity is arbitrary and can be expressed as a polynomial
- $\lambda(y) = \Sigma_{k=0}^{\infty} a_k y^k$
- In prices theorem choose g(x,y) as

$$g(x,y) = x\,\lambda(y)$$

$$\frac{\partial \mathrm{E}(\mathrm{x}\lambda)}{\partial \rho} = \frac{\partial R_{\lambda x}}{\partial \rho}$$

$$\frac{\partial \mathrm{E}(\mathrm{x}\lambda)}{\partial \rho} = \frac{\partial R_{\lambda x}}{\partial Ryx}$$

$$E\left[\frac{\partial^{2n} g(x,y)}{\partial x^n \partial y^n}\right] = \frac{\partial^2\left[x\left(\sum_{k=0}^{\infty} a_k y^k\right)\right]}{\partial x \partial y} = E\left[\sum_{k=1}^{\infty} a_k y^{k-1}\right]$$

$$= \sum_{k=1}^{\infty} a_k E\left[y^{k-1}\right]$$

=constant

X and y is jointly gaussian

$$\frac{\partial R_{\lambda x}}{\partial Ryx} = \mathbb{C}$$

Week 06 lecture 27

- Estimate the rate response of neuron

- Let r(t() be the rates at every time t , s(t) represent the stimulus train , then r(t) can be represented as below

  The estimate of rate is

$$\begin{bmatrix} S_t(1) \\ S_t(2) \\ S_t(M) \end{bmatrix} = \overrightarrow{s_t}$$

$$\overrightarrow{h^t} \cdot \overrightarrow{S_t}$$

$$r[t] = (h_1, h_2, h_3 \ldots h_M) \begin{pmatrix} \\ S_t(1) \\ S_t(2) \\ S_t(M) \end{pmatrix}$$

$$\hat{r}(t) = \overrightarrow{h_t}\, \overrightarrow{S_t}$$

  Mean square error of the linear model

  = (Estimated rate - actual rate)^2 and average it over all the time points

Note:- r is a scalar , so $\vec{S} r$ and $s^T r$ are the same.Its the cross correlation of S and r.

$$< \vec{S}\, \overrightarrow{S^T} > = C_{ss}$$

$$< \vec{S}\, r > = C_{rs}$$

$$E = \langle (\hat{r} - r)^{\wedge}2 \rangle\_t$$

$$= \langle ((h^{\wedge}T)^{\rightarrow} S^{\rightarrow} - r)^{\wedge}2 \rangle$$

$$= \langle ((h^{\wedge}T)^{\rightarrow} S^{\rightarrow} - r)\, ((h^{\wedge}T)^{\rightarrow} S^{\rightarrow} - r)^{\wedge}T \rangle$$

$$= \langle ((h^{\wedge}T)^{\rightarrow} S^{\rightarrow} - r)((h)^{\rightarrow}\ (S^{\wedge}T)^{\rightarrow} - r) \rangle$$

$$= (\langle (h^{\wedge}T)^{\rightarrow} (S)^{\rightarrow} (h)^{\rightarrow}\ (S^{\wedge}T)^{\rightarrow}) \rangle - \langle (r(h^{\wedge}T)^{\rightarrow} S^{\rightarrow}) \rangle - \langle (r(h)^{\rightarrow}\ (S^{\wedge}T)^{\rightarrow}) \rangle + \langle r^{\wedge}2 \rangle$$

Replacing C_ss and C_rs

$$= (h^{\wedge}T)^{\rightarrow}\ C\_ss\,(h)^{\rightarrow} - 2\,C\_rs\,(h)^{\rightarrow} + \langle r^{\wedge}2 \rangle$$

Minimizing the error

$$dE/\,dh = 2\,C\_ss\,(h)^{\rightarrow} - 2\,C\_rs\,(h)^{\rightarrow} = 0$$

$$\Rightarrow\ (h)^{\rightarrow} = (C_{ss})^{-1} C_{rs}$$

Reconstruction of stimulus from rates

-To estimate the stimulus at a point , we need the future set of rates

- let M bins be present in spikes train after time point ,t

- the Mth bin is at timer point (t+M)

Then

The rate vector can be represented as

$$\overrightarrow{r_{t+M}} = \begin{bmatrix} r_1 \\ r_2 \\ . \\ . \\ r_M \end{bmatrix}$$

$$\hat{S}[t] = \overrightarrow{h_t}\,\overrightarrow{r_{t+M}}$$

To ensure that the estimated stimulus is also zero mean

$\hat{S}[t] = \overrightarrow{h_t}\,(\vec{r} - \bar{r})$

To find the h value, if we replace $(\vec{r} - \bar{r}) = d$

- ⇨ Then, $\vec{h} = C_{dd}^{-1} C_{ds}$

So, after observing M bins of rate, we get first point of stimulus and on proceeding the same way, we can reconstruct the entire stimulus

Week 06 Lecture 28,29

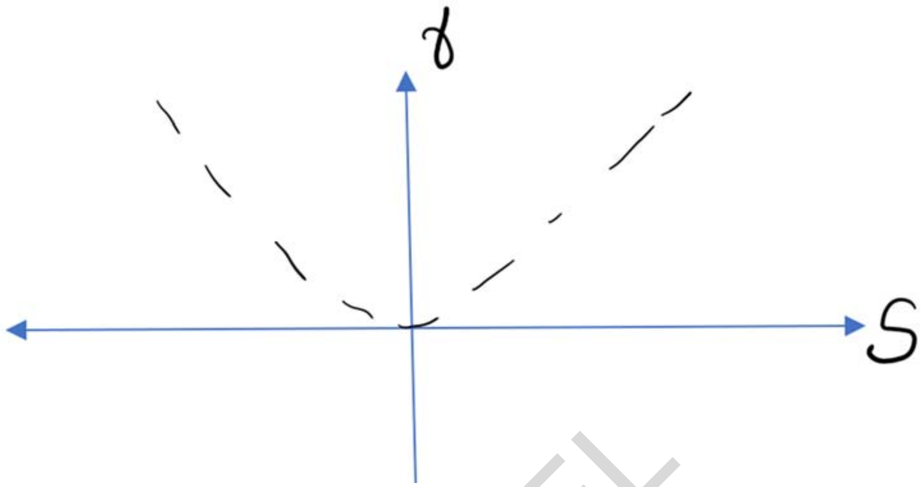
From Thomas and Young - Elements of Information Theory, refer the following pages:
Chapter 2 Entropy, Relative Entropy and Mutual Information (Page 12)
2.1 Entropy
2.2 Joint Entropy and Conditional Entropy
2.3 Relative Entropy and Mutual Information
2.4 Relationship between entropy and mutual information
2.8 Data Processing Inequality

On page 40, There is a summary in a box titled "Summary of Chapter 2". Please refer to this.

You can also check some notes under "Information theory" section in this link

Week 06 Lecture 30

If the response and stimulus have a quadratic relationship , r=as^2+n
n~(0,$\sigma^{2)}$)



Then the C_{rs} = E[rs] - E[r] E[S]

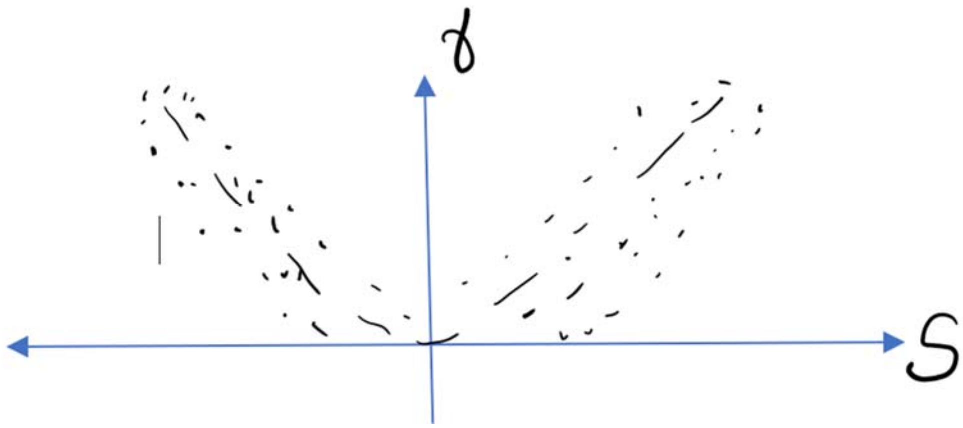
= $E[(as^2 + n)s]$

= $E[as^3]\ + E[ns]$

= $a\ E[s^3]\ +\ E[n]\ E[s]$

E[s] ~ 0 ; when its zero mean

When we consider multiple stimulus and plot responses ,



– Refer https://www.princeton.edu/~wbialek/rome/refs/sharpee+al_04.pdf

- Consider a stimulus space $S_1, S_2, \ldots\ldots\ldots S_n$ where there is a relevant subspace
- Each stimulus is a n dimensional parameter in a stimulus space .
- Refer to the auditory and visual stimulus space in video lecture 15:00
- How to determine the subspace that provide the responses relevant to a particular stimuli ?
- The probability of response given stimulus $\vec{S}, P(spikes|\vec{S})$

The average information carried by the arrival time of one spike is given by Brenner, Strong, et al. (2000),

$I_{spike} = \int_s \quad ds\ P[\vec{S}|\ spike]\ log_2\ (P[\vec{S}|\ spike]\ /\ P[\vec{S}])$

where ds denotes integration over full D–dimensional stimulus space.

- Our assumption is that spikes are generated according to a projection onto a low-dimensional subspace. Therefore, to characterize relevance of a particular direction v in the stimulus space, we project all of the presented stimuli onto v and form probability distributions $P_v(x)$ and $P_v(x|spike)$ of projection values x for the a priori stimulus ensemble .
- $\overrightarrow{S}.\vec{v} = x$

- Then ,
- $I_{spike}(\overrightarrow{v)} = \int_x \quad dx\, P_v(x|spike)\, log_2\, (P_v(x|spike)$ / $P_v(x))$

—---