Exercise I 1/29/2022

Section 2.1

- derive the eq. for MLE for $\lambda_j$

$j$: grating orientation in the experiment

$$\log p(y_{j,1:N_j} \mid x_j, \lambda_j) =$$

$$\log(\lambda_j \Delta) \sum_{n=1}^{N} y_{j,n} - N_j \lambda_j \Delta - \sum_{n=1}^{N_j} \log(y_{j,n}!)$$

goal: take derivative of log-likelihood w.r.t $\lambda_j$, set equal to $\emptyset$, solve for $\lambda_j$

① derivative w.r.t $\lambda_j$:

$$\frac{d}{d\lambda_j} \log p(\ldots) = \left(\frac{1}{\lambda_j \Delta} \cdot \Delta\right) \sum_{n=1}^{N} y_{j,n} - N_j \Delta - 0$$

$$= \frac{1}{\lambda_j} \sum_{n=1}^{N} y_{j,n} - N_j \Delta$$

② set derivative of log-likelihood to 0 & solve for $\lambda_j$

$$0 = \frac{1}{\lambda_j} \sum_{n=1}^{N} y_{j,n} - N_j \Delta$$

$$N_j \Delta = \frac{1}{\lambda_j} \sum_{n=1}^{N} y_{j,n}$$

$$\boxed{\lambda_j = \frac{1}{N_j \Delta} \sum_{n=1}^{N} y_{j,n}}$$

# neurons : 6     # orientations $j$ : 16     # trials : 40     binwidth $\Delta$ : 0.5

'spikes' : [neuron , orientation , trial ]

x : orientation     y : summed spike counts

Gaussian

$$\hat{\mu}_j = \frac{1}{N_j \Delta} \sum_{n=1}^{N_j} y_{j,n} \qquad \hat{\sigma}_j = \frac{1}{N_j \Delta} \sum_{n=1}^{N_j} (y_{jn} - \hat{\mu}_j \Delta)^2$$

Pois.

$$\hat{\lambda}_j = \frac{1}{N_j \Delta} \sum_{n=1}^{N_j} y_{jn}$$

---

Section 2.3

Gaussian log-like : $-\frac{N_j}{2} \log \sigma_j^2 - \frac{1}{2\sigma_j^2 \Delta} \sum_{n=1}^{N_j} (y_{jn} - \mu_j \Delta)^2 + c$

Poisson log-like : $\log(\lambda_j \Delta) \sum_{n=1}^{N_j} y_{jn} - N_j \lambda_j \Delta - \sum_{n=1}^{N_j} \log(y_{j,n}!)$

Neuro 316QC – Session 2

2/2/2022

speclet:
* t-dist for hyperprior distribution

main topics: Gaussian distribution, linear models

1. Univariate & Multivariate Gaussian distributions

A. mean & variance

$x = az$
$E[x] = aE[z]$
$var(x) = a^2 \cdot var(z)$

$x = az + b$
$E[x] = aE[z] + b$
$var(x) = a^2 E[z]$

$x = a_1 z + b + a_2 y$
$E[x] = a_1 E[z] + b + a_2 E[y]$
$var(x) = a_1^2 var(z) + a_2^2 var(y) + 2a_1 a_2 cov(x, y)$

2. Univariate Gaussian

- log prob. of Gaussian drops off quadratically from the mean
  - "light tails" ⇒ sensitive to data points far from the mean
- linear transformations of a Gaussian results in a Gaussian again
- Eigenvectors & values
  - eigenvectors: direction
  - eigen values: ~~amount~~ magnitude in the direction of the eigenvector

4. Probabalistic operations (marginalization & conditioning)

a. marginalization
- see explination in slides → easy for a Gaussian
- if marginalize out $x_i$, then just remove it from mean vector & its column & row from the covariance matrix

b. conditioning
- conditioning on one dimension results in another Gaussian

albertchen@g.harvard.edu

2. Linear Models

- to be a linear model, only need to be linear in model weights, not x
  - $y = w^T x + \eta \rightarrow y = w^T \phi(x) + \eta$

    $\eta$: normal noise

    ↳ $\phi(x)$: some function of x can be non-linear

3. Bayesian model comparison

- general idea: "model as another variable to condition on"
- assume two models of diff. structure: $M_1, M_2$
  - priors: $p(w \mid M_j)$
  - data likelihood: $p(X \mid w, M_j)$
  - posterior: $p(M_j \mid X) \propto p(X \mid M_j)\, p(M_j)$

    ↳ model evidence: $p(X \mid M_j) = \int p(X \mid w, M_j)\, p(w \mid M_j)\, dw$
- penalizes complexity of model automatically
  - more complex model can "explain" more data
  - $\Rightarrow$ weaker $p(X \mid M_j)$ because must sum to 1 & is more spread out over more values of X

To-Do:

- read paper
- prepare presentation for paper — can do more than one slide
- read "statistical methods sections" in Canvas