Statistical models of visual neurons

Anna Sotnikova
Applied Mathematics and Statistics, and Scientific Computation program

Advisor: Dr. Daniel A. Butts
Department of Biology
Project Background

Project Background

General question: which algorithm did your brain use to analyze the picture?
Project Background

David Hubel and Torsen Wisel “Cat experiment”, 1958

**Project Background**

Stimulus \( s(t) \) => ? \( \rightarrow \) Response \( r(t) \)

**Question**: What is the mathematical form of \( f(s) \)?

1. Linear
2. Quadratic
3. Cascade
Project Goals

• Implement 5 models:
  1. Linear models
     * Spike Triggered Average (STA)
     * Generalized Linear Model (GLM)
  2. Quadratic models
     * Spike Triggered Covariance (STC)
     * Generalized Quadratic Model (GQM)
  3. Cascade model
     * Nonlinear Input Model (NIM)

• Test models on 2 data sets:
  1. Retina ganglion cells (RGC) and primary visual cortex (V1) synthetic data
  2. Lateral geniculate body experimental data

Do not require parameters fitting

Require parameters fitting
Spike Triggered Average

Spike-triggered average (STA)

Experiment

Response

Analysis

STA & STC formulas

STA - Spike Triggered Average

\[ \hat{S} = \frac{1}{N} \sum_{i=1}^{N} s_i \]

STC - Spike Triggered Covariance

\[ C = \frac{1}{N - 1} \sum_{i=1}^{N} (s_i - \hat{s}) \cdot (s_i - \hat{s})^T \]

\( N \) - total number of spikes per experiment time

\( s_i \) - stimulus block preceding the \( i \)th spike

Reference [5]
Generalized Linear Model

\[ r(t) = F\left( s(t) \cdot k + h_s \cdot R(t) + b \right) \]

Notations

- \( r(t) \) - firing rate
- \( F \) - spiking non-linearity \( F(g) = \log(1 + e^g) \)
- \( b \) - shift spiking non-linearity
- \( k \) - filter
- \( h_s \) - spike history term
- \( R(t) \) - history of observed spike train

Reference [2]
Generalized Quadratic Model

\[ r(t) = F \left( k_L s + s^\top C s \right) = F \left( k_L s + \sum_{i=1}^{M} w_i (k_i \cdot s)^2 \right) \]

Notations

- \( r(t) \) - firing rate
- \( F \) - spiking non-linearity
- \( k_L \) - linear filter
- \( C \) - quadratic component
- \( k_i \) - squared filters
- \( w_i \) - +1 or -1

Reference [1]
Nonlinear Input Model

\[ r(t) = F\left( X(t) \cdot h + \sum_{i=1} w_i f_i(k_i \cdot s(t)) \right) \]

Notations

- \( r(t) \) - firing rate
- \( F \) - spiking non-linearity
- \( X(t) \cdot h \) - additional observables
- \( f_i \) - ‘upstream nonlinearities’
- \( k_i \) - squared filters
- \( w_i \) - +1 or -1

Reference [1]
Parameters estimation

1. Estimate filters and non-linear function $f_i$
   - fix $f_i$ and estimate filters

   $$f(x) = \begin{cases} 
   0 & \text{if } x \leq 0 \\
   x & \text{otherwise} 
   \end{cases}$$

   $$LL = \sum_t \left( R_{obs}(t) \log(r(t)) - r(t) \right)$$

   $$\frac{\partial LL}{\partial k_{i,m}} = \sum_t \left( \frac{R_{obs}(t)}{r(t)} - 1 \right) F'[G(t)] w_i f'_i(g_i(t)) s_m(t),$$

   $$G(t) = \sum_{i=1} g_i(t) = \sum_{i=1} w_i f_i(k_i \cdot s(t))$$

Example for NIM
Parameters estimation

- fix $k_i$ and estimate $f_i$ nonparametrically

$$f_i(g) = \sum_j a_{ij} \phi_j(g)$$

$$\varphi_k(x) = \begin{cases} 
\frac{x-x_{k-1}}{x_k-x_{k-1}} & \text{if } x \in [x_{k-1}, x_k] \\
\frac{x_{k+1}-x}{x_{k+1}-x_k} & \text{if } x \in [x_k, x_{k+1}] \\
0 & \text{otherwise}
\end{cases}$$
Parameters estimation

2. Find how many filters do we need to work with the data and $w_i$ sign.
Validation and Testing

- Validate program routines on synthetic data
  - Program should produce known output for retinal ganglion cells

Reference [1]
Validation and Testing

• Use cross-validation for LGN data set

• Test part of the data for predicting the output and comparing it with the real one

Data set description

Data set contains following elements:

- stimulus
- time interval of the stimulus update
- spike times (in units of seconds)
Implementation

Hardware

• MacBook Air, 1.4 GHz Intel Core i5, 4 GB 1600 MHz DDR3

Software

• Matlab_R2015b
Project schedule

October - mid November

- Implement STA and STC models
- Test models on synthetic data set and validate models on LGN data set

November - December

- Implement GLM
- Test model on synthetic data set and validate model on LGN data set

January - March

- Implement GQM and NIM
- Test models on synthetic data set and validate models on LGN data set

April - May

- Collect results and prepare final report
Deliverables

- Matlab code for all 5 models
- List of models’ parameters
- Reports and presentations
References


