Homework 8: Generalized Likelihood Ratio Test

Instructor: Daniel L. Pimentel-Alarcón   Due 11/15/2017

Problem 8.1 (Neural activity). In this problem you will apply many of the machine learning concepts we have looked at in this course in a brain mapping tool called *functional magnetic resonance imaging* (fMRI). In fMRI, a series of MRI brain images are collected over time. Because oxygenated and deoxygenated hemoglobin have slightly different magnetic characteristics, variations in the MRI intensity indicate areas of the brain with increased blood flow and hence neural activity.

The central task in fMRI is to reliably detect neural activity at different spatial locations (pixels) in the brain to produce an *activation map*, i.e., a binary image of *active* and *non-active* brain regions. However, the data are noisy and the variation in intensity due to activation is very subtle, which often makes this problem challenging.

The file `thumb_data.mat` contains the following data:

- A *reference* vector $\boldsymbol{\mu} \in \mathbb{R}^D$ with $D = 122$ oximetry measurements taken at the tendon of the thumb over time, as the thumb moves up and down.

- A sequence of 64×64 axial brain MRIs $\mathbf{X}_1, \ldots, \mathbf{X}_D$ taken over time, as the thumb moves up and down. These are stored in a 3-dimensional data array `X`. In Matlab code, $\mathbf{X}_t = $ `X(:,:,t)`, with $t = 1, \ldots, D$.

The objective of this problem is to develop your own procedure to produce an activation map indicating the brain regions related to the movement of a thumb. In other words, you want to determine which pixels are correlated to the movement of the thumb.

To this end, we can model the time-series of each pixel $\mathbf{x}_{ij} \in \mathbb{R}^D$ as

$$
\begin{aligned}
H_0 : \ & \boldsymbol{x}_{ij} = \boldsymbol{\epsilon}_{ij} & & \text{if pixel (i,j) is inactive,} \\
H_1 : \ & \boldsymbol{x}_{ij} = \boldsymbol{\mu} + \boldsymbol{\epsilon}_{ij} & & \text{if pixel (i,j) is active,}
\end{aligned}
\tag{8.1}
$$

where $i, j = 1, \ldots, 64$ and $\boldsymbol{\epsilon}_{ij} \in \mathbb{R}^D$ represents the noise at pixel (i,j). In Matlab code, $\mathbf{x}_{ij} = $ `X(i,j,:)`.
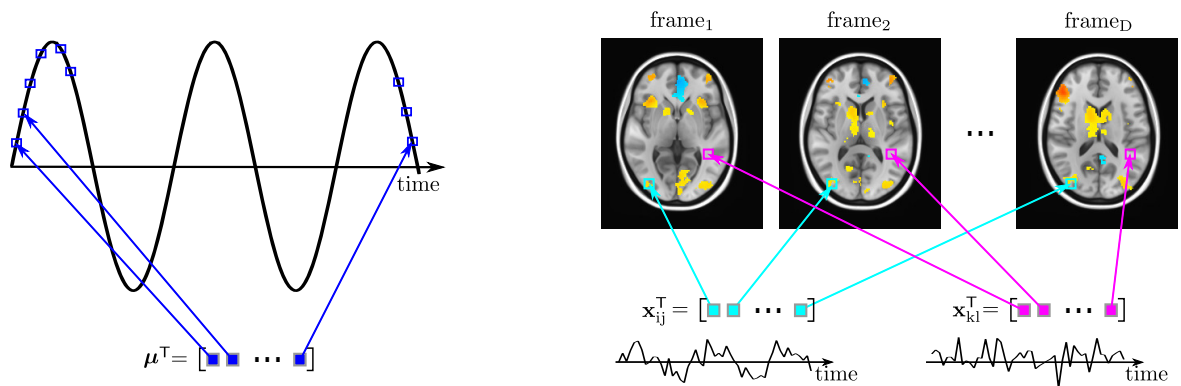

Figure 8.1: **Left:** Ideal signal of the thumb $\boldsymbol{\mu} \in \mathbb{R}^D$, here depicted as a sinusoid with the periodicity of the thumb movement. **Right:** Each pixel produces one signal vector $\boldsymbol{x} \in \mathbb{R}^D$ containing the brain measurements in that pixel over time. Some pixels may show neural activity correlated with the signal of the thumb. We want to find such voxels. See Problem 8.1.

You can get started with the (very crude) program `thumb_test.m` below that loads the data and forms a rough activation map using a simple threshold and a correlation test.

In developing your procedure, you should address/consider the following issues:

Noise in fMRI

(a) Single pixel time-series analysis.

- Is there a slowly-varying drift in the baseline? If so, how would you modify the hypotheses in (8.1) to account for this?
- Is the noise white (uncorrelated)?
- Is the noise Gaussian?

(b) Correlation between spatially neighboring pixel time-series analysis.

- Is the noise in different pixels (spatial locations) uncorrelated?
- Is the noise jointly Gaussian?

fMRI data

(c) Are the data vectors $\mathbf{x}_{ij}$ at different spatial locations similar to the reference vector $\boldsymbol{\mu}$?

(d) Are some pixels negatively correlated with the reference vector?

(e) Is there a possible delay in the timing of the data at different spatial locations?

(f) Are there correlations between the data at neighboring locations? For example, if a pixel at one location appears to be active, is it more or less likely that its neighbors are active?

GLRT for fMRI

Based on your analysis, derive a testing procedure for fMRI using the *generalized likelihood ratio test* (GLRT).

(g) The GLRT should be invariant to uncertainties in noise and baseline drift characteristics.

(h) Investigate the performance of your GLRT. Determine the distribution of the test (if possible), and summarize performance with ROC curves.

(i) Discuss advantages and disadvantages of your method, and suggest avenues for future work in fMRI.

(j) Consider the multiple testing issue here? How should the threshold of the test be adjusted in order to guarantee a desired family-wise error level, say $\alpha = 0.05$?

(k) How many false alarms might you expect without controlling the FWER?

`thumb_test.m`

```matlab
1   % ===== Code to generate a crude brain activation map using a simple =====
2   % ===== threshold and a correlation test.
3
4   clear all; close all; clc;
5   load thumb_data;
6
7   % ===== Plot reference vector =====
8   figure(1);
9   plot(mu);
10  title('Reference vector \mu');
11
12  % ===== Compute correlation of all pixels with reference =====
13  C = zeros(64,64);
14  E_mu = mean(mu);
15  for i=1:64,
16      for j=1:64,
17          x_ij = X(i,j,:);
18          x_ij = reshape(x_ij,[],1,1);
19          C(i,j) = (x_ij-mean(x_ij))' * (mu-E_mu);
20      end
21  end
22
23  % ===== Set a threshold at 35% of maximum to produce activation map =====
24  th = .35 * max(max(C));
25  Act_Map = C>th;
26
27  % ===== Display correlation image and activation Map (in red) =====
28  figure(2);
29  imagesc(C);
30  axis('square');
31  colormap('gray');
32  hold on;
33  [j,i] = find(Act_Map);
34  plot(i,j,'rs');
35  title('Activation Map');
```
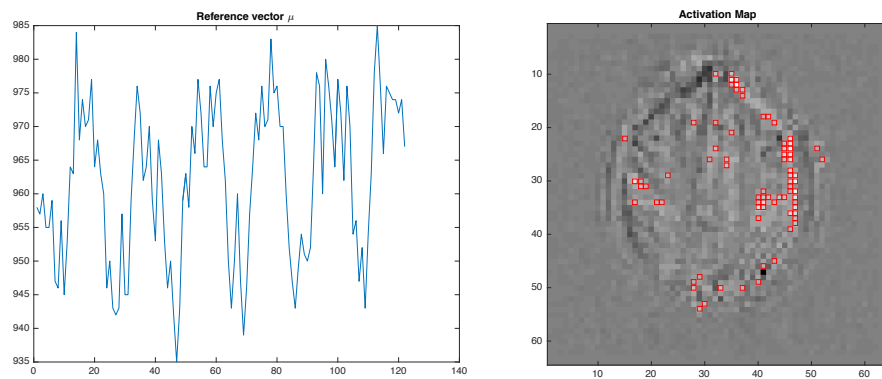

Figure 8.2: Output of the program `thumb_test.m` from Problem 8.1. **Left:** Reference vector $\boldsymbol{\mu} \in \mathbb{R}^D$ produced by the thumb movement. **Right:** Rough brain activation map produced by a simple threshold and a correlation test.