

Spongebob Squarepants

asophonsri1

July 2019

Aims

Focus on aspects of BOLD that is relevant to us

1. Find the individual contributions of O_2 and CO_2 to bold
2. Introduce fourier transform as a method to clean physiological noise from bold
3. Explore the use of gradient descent

7/19/2019

Logistic Regression

We introduce the idea of a continuous logistic function used to regress to logistically behaving phenomena. Some such phenomena include bacterial growth in an environment with limited carrying capacity, the advancement of computer technology, but more importantly, the response curves for End-Tidal gasses recorded during a CVR-FMRI test.

The logistic curve:

$$\zeta = \frac{M}{1 + p^{-\alpha t}} + \psi$$

is uniquely characterized by asymptotic behaviour as $t \rightarrow \pm \infty$

However, regressing to the logistic curve poses a problem; the logistic curve innately relies on 4 independent variables: p , α , M , and ψ . Therefore, in order to optimize the regression process, we propose using gradient descent as an iterative logistic regressor. As a cost function, ordinary least squares will have to do:

$$\text{cost} = \left(\sum_{t=0}^N [\text{model}(t) - \text{observed}(t)]^2 / N \right)$$

The gradient of this cost returns a vector of 4 variables:

$$\vec{\delta} = \nabla(\text{cost}) = \left\langle \frac{\delta}{\delta p} \zeta, \frac{\delta}{\delta \alpha} \zeta, \frac{\delta}{\delta M} \zeta, \frac{\delta}{\delta \psi} \zeta \right\rangle$$

where:

$$\frac{\delta}{\delta p} \zeta = \frac{2}{N} \left(\frac{M}{1+p^{-\alpha T}} + \psi - Y \right) \left(\frac{M \alpha T p^{-\alpha T-1}}{(1+p^{-\alpha T})^2} \right) \quad (1)$$

$$\frac{\delta}{\delta \alpha} \zeta = \frac{2}{N} \left(\frac{M}{1+p^{-\alpha T}} + \psi - Y \right) \left(\frac{M \ln(p) T p^{-\alpha T}}{(1+p^{-\alpha T})^2} \right) \quad (2)$$

$$\frac{\delta}{\delta M} \zeta = \frac{2}{N} \left(\frac{M}{1+p^{-\alpha T}} + \psi - Y \right) \left(\frac{1}{1+p^{-\alpha T}} \right) \quad (3)$$

$$\frac{\delta}{\delta \psi} \zeta = \frac{2}{N} \left(\frac{M}{1+p^{-\alpha T}} + \psi - Y \right) \quad (4)$$

under the case that:

$$T = \sum_{t=0}^N t$$

$$Y = \sum_{t=0}^N y(t)$$

We can now use this cost gradient in regression.

$$\vec{\Omega}_{i+1} = \vec{\Omega}_i + \varepsilon \vec{\delta}$$

where $\varepsilon \in \mathbb{R}$ is sufficiently small and $\vec{\Omega}$ are parameter tuples

7/22/2019

Logistic Regression Corrected

The mistake from above is assuming that the exponential base is not euler's constant: e Additionally, we should consider fitting an initial t_0 as well. This should result in a slightly different gradient function, but should stay more or less the same

Our corrected logistic regressor is:

$$\zeta = \frac{M}{1 + e^{-\alpha(t-\phi)}} + \psi$$

It follows that the basis of the cost gradient operation is:

$$\left\langle \frac{\delta}{\delta M} \zeta, \frac{\delta}{\delta \alpha} \zeta, \frac{\delta}{\delta \phi} \zeta, \frac{\delta}{\delta \psi} \zeta \right\rangle$$

And therefore, the gradient includes the terms:

$$\begin{aligned}
\frac{\delta}{\delta M} \zeta &= \sum_{t=0}^T 2 \left(\frac{M}{1 + e^{-\alpha(t-\phi)}} + \psi - y \right) \frac{1}{1 + e^{-\alpha(t-\phi)}} \\
\frac{\delta}{\delta \alpha} \zeta &= \sum_{t=0}^T 2 \left(\frac{M}{1 + e^{-\alpha(t-\phi)}} + \psi - y \right) \frac{M}{(1 + e^{-\alpha(t-\phi)})^2} \left[(t - \phi) e^{-\alpha(t-\phi)} \right] \\
\frac{\delta}{\delta \phi} \zeta &= \sum_{t=0}^T 2 \left(\frac{M}{1 + e^{-\alpha(t-\phi)}} + \psi - y \right) \frac{-\alpha M e^{-\alpha(t-\phi)}}{(1 + e^{-\alpha(t-\phi)})^2} \\
\frac{\delta}{\delta \psi} \zeta &= \sum_{t=0}^T 2 \left(\frac{M}{1 + e^{-\alpha(t-\phi)}} + \psi - y \right)
\end{aligned}$$

Fourier Transform

Uniform vs Non-uniform

Discrete fourier transforms are functions that map time-varying signal (generally periodic) to a histogram indicating the distribution of frequencies. Discrete Fourier transforms are found using:

$$\begin{aligned}
\hat{\chi} &= \sum_{n=0}^{N-1} \chi_n \cdot e^{-\frac{i2\pi}{N}kn} \\
&= \sum_{n=0}^{N-1} \chi_n \cdot \left[\cos\left(\frac{2\pi kn}{N}\right) - i \cdot \sin\left(\frac{2\pi kn}{N}\right) \right]
\end{aligned}$$

In which we can abstract:

$$\frac{k}{N} \rightarrow \text{frequencies} \quad (5)$$

$$n \rightarrow \text{time} \quad (6)$$

However, it is not the case that signals are sampled periodically; luckily, we (someone else) has abstracted the idea of the discrete fourier transform to the following:

$$\begin{aligned}
\hat{\chi} &= \sum_{n=0}^{N-1} \chi_n \cdot e^{-2\pi i p_n \omega_k} \\
&= \sum_{n=0}^{N-1} \chi_n \cdot \left[\cos(2\pi p_n \omega_k) - i \cdot \sin(2\pi p_n \omega_k) \right]
\end{aligned}$$

In which we have abstracted uniform frequency to a mapping of frequencies: ω_k and a uniform timestep to a mapping of timesteps: p_n

Both inverse transforms are defined by replacing:

$$\hat{\chi} \text{ with } \chi_n \text{ as well as } (-i) \text{ with } (i)$$

Spectral Leakage

Spectral Leakage is the phenomena where the application of a time-variant function introduces new frequencies into the frequency distribution. In this case, the rectangular windowing of the transform

introduces some frequency scalloping: the target frequency we want to extract is clouded by similar frequency. We want to explore different windowing functions to optimize frequency extraction. Notably, we are experimenting with Blackman and Hamming windows.

Read Window Advantages (email Jimmy)

Notes on sampling rate of GE scanner

- GE respiratory scan rate is 40 ms, we want minute-frequency so we divide .04 by 60
- BOLD data is measured in fractions of minutes no division is necessary to convert to frequency domain
- Pulse Ox measures every 10 ms and ECG measures every 1 ms

7/23/2019

Fourier Transform

To recap from yesterday, the inputs to the fourier transform are indeed uniform. The current problems faced are:

1. Summation (subtraction) of the Power Spectra yields negative progress; the resultant power spectrum contains more noise
2. True (non-truncated) algebra on the Power Spectra is not currently possible due to different numpy array sizes

In addition to the problems, we have observed several phenomena related to fourier transforms:

1. The fourier transform is symmetric across the perpendicular bisector of the frequency domain.
2. Low pass filters preserve larger features.

That being said, we have observed good preservation of feature with a low-pass filter filtering frequencies $> \approx .0583$ hz We think the resulting noise is caused by Heart rhythm

We suggest the following method:

1. smooth the respiratory power spectrum using a mirrored-symmetric boundary condition IIR filter
 - the resulting spectra should be approximately gaussian
2. calculate the range of acceptable confidence. We will define acceptability later.
3. zero the calculated range in the power spectra of the CO₂/O₂ signal
 - remember that operations must be performed symmetrically

It is possible that this method yields little benefit.

Spectral Leakage

Non-rectangular windows distort uneventful data past usefulness. We should formulate an algorithm to choose whether or not to use a window, and we should experiment to see whether a window is necessary.

Comparison with Feat

TODO

1. consider piece-wise fitting: the entire series not of the same model
2. consider temporal derivative piece-wising
3. consider filtering breathing rhythm from data
4. draw graphs
5. insert data drawings: time series data as well as gradient descent data
6. test atom