BME I5100: Biomedical Signal Processing
(Was Non-linear signal processing in biomedicine)

Introduction

Lucas C. Parra
Biomedical Engineering Department
City College of New York
Biomedical Signal Processing - Content

We will cover basic principles of signals processing. We will emphasize examples and focus on electrical signals generated by the biological systems (biopotentials).

We will introduce concepts from:
• filter theory
• statistical processes
• pattern recognition
• information theory
• probabilistic modeling
• Neurophysiology
• MATLAB

Prerequisite:
linear algebra, some programing language, complex variables.

Literature
Eugene N. Bruce, Biomedical Signal Processing and Signal Modeling, John Wiley & Sons, 2000
Schedule

**Week 1: Introduction**
Linear, stationary, normal - the stuff biology is not made of.

**Week 1-4: Linear systems** (mostly discrete time)
Impulse response
Moving Average and Auto Regressive filters
Convolution
Discrete Fourier transform and z-transform

**Week 5-7: Random variables and stochastic processes**
Random variables
Multivariate distributions
Statistical independence

**Week 9-14: Examples of biomedical signal processing**
Probabilistic estimation
Linear discriminants - detection of motor activity from MEG
Harmonic analysis - estimation of heart rate in Speech
Auto-regressive model - estimation of the spectrum of 'thoughts' in EEG
Independent components analysis - analysis of MEG signals
Klaman Filtering – motion estimation
Schedule

**Week 1: Introduction**
Linear, stationary, normal - the stuff biology is **not** made of.

**Week 1-4: Linear systems** (mostly discrete time)
Impulse response
Moving Average and Auto Regressive filters
Convolution
Discrete Fourier transform and z-transform

**Week 5-7: Random variables and stochastic processes**
Random variables
Multivariate distributions
Statistical independence

**Week 8: Electrophysiology**
Origin and interpretation of Biopotentials

**Week 9-14: Examples of biomedical signal processing**
Probabilistic estimation
Linear discriminants - **detection** of motor activity from MEG
Harmonic analysis - **estimation** of hart rate in Speech
Auto-regressive model - **estimation** of the spectrum of 'thoughts' in EEG
Independent components analysis - **analysis** of MEG signals
Biomedical Signal Processing and Signal Modeling

**Biomedical Signal Processing** - Signal processing and statistical modeling methods useful when analyzing biomedical signals, e.g.

- Electro and Magneto Encephalography
- Electro Myograms and Cardiograms
- Circadian rhythm in body temperature
- Spike trains
- Speech
- ...

Property of BioMed signals: non linear, non stationary, non Gaussian
Linear, Stationary, Normal - The stuff biology is not.

Linear transformation \( y = L[a] \):

\[
y(t) = L[a \, x_1(t) + b \, x_2(t)] = a \, L[x_1(t)] + b \, L[x_2(t)]
\]

Physics often calls for linear combination of signals:
- Mass, force, energy
- Concentrations in solutions.
- Electrical and magnetic fields.
- Intensity of incoherent electromagnetic radiation (X-ray, visible light, radio-waves)
- Amplitude of acoustic signal.
- ....

Lets look for example at EEG

\[
>> \text{load eeg.mat}
\]
Example: We record frontal EEG electrode $y(t)$. It will be contaminated with eye muscle activity. Assume eye muscle activity generates electrical source signal, $x_1(t)$, and some other frontal brain activity gives source, $x_2(t)$. Physics tells us that electrical potentials add up linearly:

$$y(t) = [a \ b] \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

where $[a \ b]$ represent the coupling coefficients for eye muscle and frontal activity respectively.

$$y(t) - a x_1(t) = b x_2(t)$$

Linearity is crucial because given an estimate of $x_1(t)$ and $a$ for example from an electro-oculogram (EOG) we can subtract its influence on $y(t)$:
Linear, Stationary, Normal - *The stuff biology is not.*

Unfortunately signals are often distorted by non-linearity.

Common problem is limited dynamic range.

\[
\begin{align*}
\text{plot}([-x(1,:); x(8,:); x(2,:)])' \\
\text{plot}(-x(1,:),x(8,:)) & \quad \text{plot}(-x(1,:),x(2,:))
\end{align*}
\]
Linear, Stationary, Normal - The stuff biology is not.

Note that non-linearity can often be identified even in a 1D signal by its harmonic distortions.

```
>> psd(sin(x))
>> psd(atan(sin(x)))
```

Harmonic distortion
Harmonic distortion explained ...

For example distortion of quadratic non-linearity leads to frequency doubling:

$$x(t) = \sin(\omega t)$$

$$y(t) = x^2(t) = \sin^2(\omega t) = \frac{1}{2} - \frac{1}{2} \cos(2\omega t)$$

Cubic leads triple frequencies:

$$y(t) = \sin^3(\omega t) = \frac{3}{4} \sin(\omega t) - \frac{1}{4} \sin(3\omega t)$$

General non-linearity contains all orders according to Taylor expansion:

$$y = f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} \left[ \frac{\partial^n f(x)}{\partial x^n} \right]_{x=0} x^n$$
Often 'normal' distributions are assumed, i.e. Samples are Gaussian distributed.

Important because of many nice properties of the Gaussian probability density function (pdf):

$$p(x) = \frac{1}{\sigma \sqrt{2 \pi}} \exp\left(-\frac{x^2}{2 \sigma^2}\right)$$

- Convolution of Gaussian remains Gaussian
- Product of Gaussian remains Gaussian
- Parameter $\sigma$ easy to estimate.
- Leads to least squares optimization criteria
- Sums of many random variables converges to Gaussian, e.g. Brownian motion

>> hist(randn(1000, 1))
Linera, Stationary, Normal - The stuff biology is not.

Unfortunately many natural signals are NOT Gaussian.

On the left is an example of Tongue electro-myogram (EMG):

However, if we normalize by estimate of the local standard deviation:

\[
s = \text{std}(x(i-10:i+10));
\]

\[
y(i) = x(i)/s;
\]

we obtain often something that is close to Gaussian.
The property of *heteroscedasticity* is often used in the context of financial time series. e.g. NY stock exchange index. It states that the signal is short term Gaussian with **time varying** standard deviation.
Linear, Stationary, Normal - The stuff biology is not.

There are many non-stationary signals that can be explained in first approximation as heteroscedastic. In multiple dimensions these signals are also known as spherical invariant random processes.
Linear, Stationary, Normal - *The stuff biology is not.*

- Many natural signals are **not stationary**, and **not normal**, and many systems are **not linear**.

- Analysis and signal processing is **OFTEN EASIER** if one can assume stationary, normal signals and linear systems.

- It is important to identify the nature of the signals and possibly apply preprocessing to make the assumptions simpler.

- Non-linearity may be identified simply by looking at scatter plots, or harmonic distortions if a strong oscillation is present (often 60Hz).

- Non-Gaussian properties can be identified by looking at histogram. We will use cummulants to asses 'normality' quantitatively.

All signal analysis starts by **LOOKING AT THE DATA!**
Grading

Assignment 1: Reproduce the four figures on slides 8, 9, 11, 12 from the raw data. Use the files eeg.mat and tongemg.mat.

For help on MATLAB run
   >> demo
   >> help

Useful functions
   >> lookfor
   >> whos

In particular, if you are new to matlab, please make substantial time available to run the demo programs which are a very good introduction to matlab: basic matrix operations, line plotting, matrix manipulations, 2-D plots, matlab language introduction, axis properties, graphs and matrices, and maybe some of the desktop environment demos as you see fit.
Good News - No final nor midterm exam!
Bad News - Assignments:

1. MATLAB programing
   • Turn in next week my email
   • Needs to run correctly 75% of the time for passing.
   • Needs to run perfectly 100% for the time for A+.
   • May have pop quizzes to test “undisclosed collaborations”.

2. Proofs
   • Turn in next week
   • Easy, just to exercise the notation

3. Reading
   • Understand the subject and cover gaps

4. May have pop quizzes on reading and programming assignments.
Programming Assignments

- If you copy code you will fail the course.
- Assignments due in one weeks time. Submit per email **before class**.
- Submit **single executable file** called: first_last_number.m, all lower case e.g. john_smith_3.m for John's 3rd assignment. No figures, no text files, nothing except executable code.
- Your program must load all required data. Assume that data files are in current directory. All required data will be posted on the web.
- Include 'clear all, close all' at the beginning of all programs.
- **Do not use upper case** letters for commands, e.g. Use `axis()` instead of 'AXIS()'. They may work for you but they don't work for me!
- If you had help during your work, you **MUST** name your partner. "Similar" submission are easy to spot. **Undisclosed collaborations** receive 0 credit.
- The criteria for full credit should be clear. If not, please ask in class. Do not take chances by assuming that your work is "sort of correct".