Applied Linear Algebra in Data Analysis

What is this course about?

Sivakumar Balasubramanian

Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

The world is multidimensional

- Most systems in science and engineering are multidimensional.
  Multiple degrees-of-freedom, various causes and effects, multiple inputs and outputs, multiple measurements, multiple parameters etc.

- Undergrad. engineering courses focus mostly on the analysis and design of relatively simple systems.
  - Linear single-input-single-output (SISO) system theory.
  - Fourier, Laplace, z-transforms for understanding 1D temporal or spatial phenomena.
  - Study of dynamics in mechanical engineering.

- Multidimensional problems require an additional set of mathematical tools and a different style of thinking.

The world is also uncertain

- Noise and uncertainty are inherent in the real world.
- All measurements are corrupted by unwanted sources and the outcomes of all our actions are also influenced by a variety of unmeasured, unknown factors.
- There is often some structure to this uncertainty, which can be exploited.
- Dealing with uncertainity too requires an additional set of mathematical and statistical tools.

Three major topics are covered in this course

Linear Algebra

- First step in high dimensional thinking.

Probability theory

- Principled approach to navigate uncertainity.

Linear Regression

- A fundamental tool for modeling relationships between variables for the purpose of statistical inference, prediction, and causal analysis.

Linear Algebra

Thomas Garrity say in his book *All the Mathematics You Missed*,

> *". . . Though a bit of an exaggeration, it can be said that a mathematical problem can be solved only if it can be reduced to a calculation in linear algebra."*

Algebra (Arabic: *al-jabr* 'reunion of broken parts') is the study of variables and the rules for manipulating these variables in formulas (Source: Wikipedia).

Linear Algebra

- Linear algebra is the study of linear system of equations.

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= y_1 \\ &\vdots \\ a_{m1}x_1 + a_{32}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= y_m \end{aligned}\right\} \longrightarrow \mathbf{Ax} = \mathbf{y}$$

  These are useful in a surprisingly wide range of real world problems!

$$\begin{aligned} x_1, x_2, \ldots x_n &: \text{inputs/parameters of the problem} \\ y_1, y_2 \ldots y_m &: \text{outputs or measurements} \\ a_{ij} &: \text{physics of the problem} \end{aligned}$$

- Linear algebra provides tools for: understanding, manipulating, and efficiently solving such problems.

Probability Theory

- Branch of mathematics dealing with probability - a measure of uncertainty of events.
- Provides the tools for representing, analysing, and manipulating uncertainity of events.
- Some essential concepts: probability distributions/densities, random variables, expectation, variance, covariance, conditional probability, Bayes rule, etc.
- We will also look at the Gaussian or Normal distribution, arguably the most important distribution in statistics.

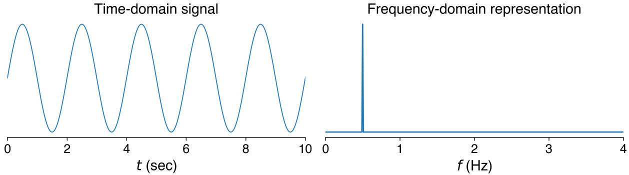
What kind of problems can we solve with these tools?

- Signal processing
- Control theory
- Robotics
- Statistics
- Machine learning
- Medicine
- Economics
- ...

Signal Processing

We often deal with temporal signal that can be visualized as a function of time. Is this the only way to represent and understand a signal?

A sinusoidal signal's time domain representation is shown below. We don't need the signal's entire time record to convey its information. We only need three numbers - the amplitude, frequency, and phase of the sinusoid. These are sufficient to reconstruct the signal. They can be obtained by Fourier transforming the signal into the frequency domain, i.e. a different perspective of the signal.



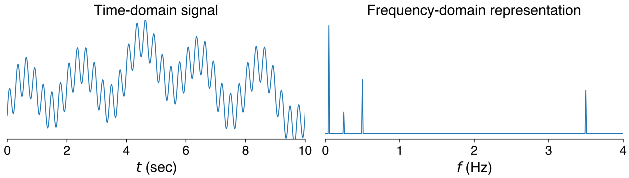
Signal Processing

- Expressing signals as a linear combinations of other signals (basis functions) is at the core of most of signal processing.

$$x(t) = c_1\phi_1(t) + c_2\phi_2(t) + \cdots + c_n\phi_n(t)$$

- $\{c_i\}_i$ is another representation of the signal $x(t)$ in terms of $\{\phi_i(t)\}_i$.
- The Fourier transform is one such representation of a signal.
- $\{\phi_i\}_i$ are called the basis functions. Different basis functions $\longrightarrow$ different types of transforms.
- Manipulating $\{c_i\}_i$ will allow us to extract or suppress different types of information from $x(t)$.

Signal Processing



What is so special about the Fourier transform? Are there other transforms that might be better suited for certain types of signals? Can I design my own transform?

Thinking of signals and images as entities residing in a high-dimensional space turns out to be very useful way to think about signals and how to process them.

Control theory

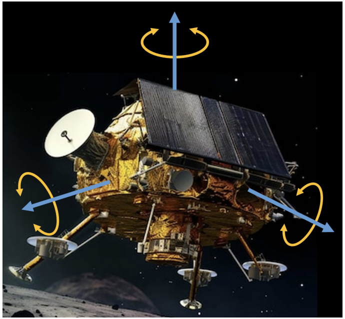
A spacecraft has 6 DOF - 3 translational

x, y, z and 3 rotational ϕ_1, ϕ_2, ϕ_3.
We need both position and velocity information to control the spacecraft.

The spacecraft has multiple sensors that measure position and velocity.

Multiple thrusters, reaction-wheels are used to control the spacecraft dynamics.

How do we get the best estimate of the spacecraft position and velocity? Which and by how much do we activate the different thrusters and reaction-wheels to control the spacecraft?



Source: http://surl.li/klulx

Robotics

Rehab Robot [Source: http://surl.li/klupd]

Surgical Robot [Source: http://surl.li/klurm]

Robotics

Kinematics and dynamics of serial robots are highly non-linear.

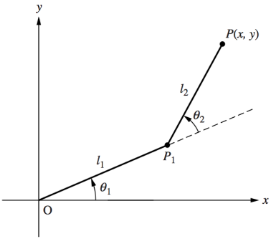
Kinematics of a 2-link planar robot

$$x = l_1 \cos(\theta_1) + l_1 \cos(\theta_1 + \theta_2)$$
$$y = l_1 \sin(\theta_1) + l_1 \sin(\theta_1 + \theta_2)$$

The differential kinematics however has is linear relationship:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \mathbf{J}(\theta_1, \theta_2) \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix}$$

$\mathbf{J}(\theta_1, \theta_2)$: Jacobian matrix



Statistics

The data that we deal with is often heterogeneous, high-dimensional, and noisy.

Core of statistics:

- Representing such data
- Evaluating the relationships between different variables
- Building and estimating models for inference and decision making

Probability theory, Calculus, Linear algebra, and Optimization are employed for solving these problems efficiently.

Linear Regression Models

A standard multivariate linear regression model.

$$y_j = \beta_0 + \beta_1 x_{j1} + \beta_2 x_{j2} + \cdots + \beta_n x_{jn} + \epsilon_j \longrightarrow \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

$$\min_{\boldsymbol{\beta}} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 \implies \hat{\boldsymbol{\beta}} = \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{y}$$

A model like this has several important uses:

- *Exploring association between variables*
- *Predicting the value of a variable from other variables*
- *Extrapolating beyond the sample data*
- *Causal inference*

Linear Regression

Richard McElreath, a well-known statistician, in his book *Statistical Rethinking*[2] writes about linear regression:

> *... This type of model is simple, flexible, and commonplace. Like all statistical models, it is not universally useful. But linear regression has a strong claim to being foundational, in the sense that once you learn to build and interpret linear regression models, you can more easily move on to other types of regression which are less normal.*

Some of the principles of linear regression carry over to more machine learning complex models.

- Decision trees, random forests
- Support vector machines
- Neural networks
- ...

What will you learn from this course?

- Basic linear aalgebra and its geometric view.
- Linear least squares methods.
- Basic probability theory and statistics.
- Understanding, constructing, and interpreting linear legressions models.