The philosophy of the course reflects that of the instructor, who takes a dualistic view about information, data, science and engineering. It recognizes the rich complexity of signals and phenomena we wish to identify and analyze while taking a minimalist reductionist view when choosing and applying analysis techniques.

My goal is to present material that allows you to understand and derive algorithms to a sufficient level that you could write the necessary code for their implementation. This does not mean that you always should write your own code. After all, many fine packages implement arsenals of tools that can be used (IDL, MATLAB, Mathematical, etc.). I suggest these rules of thumb:

- Avoid using a “canned” program in one of these packages unless you can write down the underlying mathematics and could derive and program the algorithm.
- “Experiment” with programs after you understand what they should do mathematically.
- Avoid experimenting with a program to try to infer or reverse engineer what it is claimed to do; this is very inefficient.
- Programs may not do what they claim to do or they may have built-in, restrictive assumptions.
- Always test code with toy examples before using in an important application.
A Minimal List of Signal Analysis Themes:

1. Frequentist vs. Bayesian methods
2. Detection and Discovery
3. Matched Filtering and Optimization
4. The Central Limit Theorem and Non-Gaussianity
5. Basis Vectors and Compact Support
6. Aliasing and the pros and cons of uniform sampling
7. Finding phase
8. Deconvolution Tricks (inverse problems)
9. Doing the forward problem to solve an inverse problem
10. Defeating the Uncertainty Principle
11. Deterministic vs. Chaotic vs. Stochastic signals
12. Comparing models and hypotheses (statistical inference)
13. Space Exploration:
   searching parameter spaces of high dimensionality
14. Analog vs. Digital: the effects of quantization in both x and y
Signal Analysis Themes:

1. **Frequentist vs. Bayesian methods:**
   Two approaches to probability translate into two broad approaches to data analysis and inference. One considers the outcomes of experiments in terms of frequency of occurrence and a hypothetical ensemble while the other ties probability to the state of knowledge before and after acquiring and analyzing data sets. The approach of the course is dualistic.

2. **Detection and Discovery:**
   Discovering new phenomena and objects are central to observational science. In astronomy, many challenges boil down to finding weak signals buried in noise. Finding signals or patterns amid clutter is another data mining problem that we will address.

3. **Matched Filtering and Optimization:**
   Matched filtering typically means fitting a noisy data set with a signal template that is identical to the ‘true’ signal, usually through a convolution method. Matched filtering optimizes the signal-to-noise ratio of a ‘test statistic.’ The notion of matched filtering can be extended to many procedures, including:
   - testing the existence of a signal in a data set (detection)
   - least-squares fitting of functions to data
   - estimating the time of arrival of a pulse (testing relativity with pulsar timing)
   - centroid frequency or velocity of a spectral line (red shifts, exoplanets)
   - template matching of predicted gravitational waveforms to gravity detector data
4. The Central Limit Theorem and Non-Gaussianity:

The factor $\sqrt{N}$ is ubiquitous in statistical modeling and analysis, as we all know from error analysis in laboratory contexts. However, it appears in many other instances, including errors on least-squares estimated parameters, power spectra, etc. and is directly related to the CLT, which describes convergence of the underlying PDF to a Gaussian $\equiv$ normal form.

— incoherent summing procedures
— coherent summing procedures

5. Basis Vectors and Compact Support:

Spectral analysis often means “analyze the power spectrum that is based on the Fourier Transform.” More generally, the goal is often to characterize measurements with the smallest number of underlying basis vectors. Fourier basis vectors (sinusoids) are appropriate in some contexts but not others.

(a) When is Fourier analysis appropriate? When not?
(b) Other bases: wavelets, spherical harmonics, etc.
(c) Principal Component Analysis: let the data determine the best basis vectors.

6. Aliasing and the pros and cons of uniform sampling:

Aliasing is the appearance (in Fourier analysis) of signal components at the wrong apparent frequencies. Counteracting it involves understanding the sampling theorem and the role of uniform sampling. In some instances, nonuniform sampling is beneficial for aliasing, but can make the analysis more difficult. Techniques for spectral analysis will be discussed for the case of non-uniform sampling.
7. Finding Phase:

An often encountered problem consists of inferring a function when only the magnitude of its Fourier transform is known. Bootstrapping the inference can be done by using additional information or by imposing conditions on the function, such as causality and positivity (phase retrieval). In some contexts, phase may be more important than amplitude.

8. Deconvolution Tricks (inverse problems):

Often a measurement $x(t)$ is the convolution of a quantity of interest $y(t)$ and a filter or smearing function $h(t)$. (Many natural phenomena can be characterized by such linear systems.) Typically the integral $x(t) = \int dt' y(t') h(t - t')$ partially destroys information about $y(t)$. Deconvolution means to estimate $y(t)$ from the measurements, $x(t)$. This can be done in approximate ways that are limited by the information-destroying aspects of $h(t)$ and also by the finite S/N of the measurements.

9. Doing the forward problem to solve an inverse problem:

Rather than attempting deconvolution, one can simulate the measurement process by using trial functions or processes $\hat{y}(t)$, passing them through the filter $h(t)$ to obtain the trial $\hat{x}(t)$, which is compared to the actual measurements $x(t)$. Thus we test models in “measurement space.” By iterating, the procedure may converge to a consistent (but usually not unique) answer. This approach can be far more robust than deconvolution. There are also instances where even $h(t)$ is not known, so one can use trial functions for $h(t)$ as part of the iteration process.
10. **Defeating the Uncertainty Principle:**

For frequency-time variables, the uncertainty principle is $\Delta \nu \Delta t \gtrsim 1$. This means that you can’t localize a signal in both time and frequency to arbitrary precision. In some instances, one can do better than what naive application of the uncertainty principle would suggest. This is called *superresolution* in spectral analysis and imaging applications.

11. **Deterministic vs. Chaotic vs. Stochastic signals:**

Death and taxes are deterministic events in that they are bound to happen. But they are also stochastic in that we don’t know by how much or when taxes may be reduced/increased or when one will die. Random number generators appear to produce stochastic output but they actually produce numbers comprising a chaotic process, which is a particular kind of deterministic process. How can we tell the difference for a measured data set? Procedures exist for testing the properties of a data set in this regard.
12. **Comparing models and hypotheses (statistical inference):**

If we don’t know the best model for a data set or phenomenon *a priori*, then somehow we need to determine it from measurements. Statistical inference involves determining the best parameters *given* a model, implying that we have some *goodness of fit metric* that we apply to determine the best values for the parameters. This notion can be extended to alternative models or even hypotheses.

- Frequentist inference
- Bayesian inference
- Incorporating prior knowledge and mathematical constraints
- Ensembles and realizations: estimation errors when only one realization of a process can be measured (e.g. extinction record over geologic time; cosmic evolution and cosmic variance).

13. **Space Exploration: searching parameter spaces of high dimensionality** Statistical inference often involves finding a best-fit, nonlinear solution in a parameter space whose dimensionality is too large to explore by “brute force.” Methods exist for exploring such spaces that adopt methods found in nature in thermodynamic or biological contexts. These include:

- Downhill simplex
- Simulated annealing
- Markov Chain Monte Carlo methods
- Genetic algorithms
- Neural networks
14. Analog vs. Digital:

Often we think about physics etc. in continuous terms while doing computer analysis necessarily with digital quantities. What are the consequences? Sometimes we exploit extreme types of quantization to develop a fast algorithm or hardware processor.
— Examples where sampling (digitization) and Fourier analysis do not commute.
— Correlation spectrometers.