15-830/630 – Problem Set #2

1. **Logistic Regression, Newton's Method, and Least Squares [10pts]**
   Recall from class that Newton's method repeats the updates

   $$\theta \leftarrow \theta - \left(\nabla_\theta^2 J(\theta)\right)^{-1} \nabla_\theta J(\theta)$$

   and that for the logistic regression objective function

   $$J(\theta) = \sum_{i=1}^{m} \log(1 + \exp(-y_i \cdot \theta^T \phi(x_i)))$$

   the gradient and Hessian are given by

   $$\begin{aligned} \nabla_\theta J(\theta) &= -\Phi^T Z y \\ \nabla_\theta^2 J(\theta) &= \Phi^T Z(I - Z)\Phi \end{aligned}$$

   where $Z \in \mathbb{R}^{m\times m}$ is a diagonal matrix with

   $$Z_{ii} = \frac{1}{1 + \exp(y_i \cdot \theta^T \phi(x_i))}$$

   (we didn't prove these in class, and you won't need to prove them for this problem).

   Suppose we start with $\theta = 0$ and perform one iteration of Newton's method for logistic regression. Give a closed form for θ after one iteration. Describe how this relates to least squares regression.

2. **Support Vector Machines [20 pts]**

   In this problem, you'll derive standard QP formulations and experiment with support vector machines for classification.

   (a) As mentioned in previous homework, the standard form quadratic program is written as

   $$\begin{aligned} \underset{z}{\text{minimize}} \quad & \frac{1}{2} z^T P z + q^T z \\ \text{subject to} \quad & Az \le b. \end{aligned}$$

   Formulate the support vector machine (Hinge loss plus regularization for the objective) as a standard form quadratic program. To write this formulation concisely, you'll want to use the fact that you can write the elementwise vector product

   $$c = a \circ b \;(c_i = a_i b_i, \; \forall i) \quad \Longleftrightarrow \quad c = \mathrm{diag}(a) b$$

where diag(a) constructs a diagonal matrix from the entries of the vector a. Hint: the formulation overall here is similar to the formulation you did for the deadband loss in Problem Set 1.

(b) **(15-830 only)** Write the kernelized version of the SVM as a quadratic program in standard form.

(c) Here you'll implement an SVM for the refrigerator classification task discussed in class. You'll use the file q2_data.txt, which is in the format:

<device power consumption> <device on duration> <label>

where label is either $+1$ or -1 depending on which of the two fridges the signal corresponds to. Load this data, and run an SVM using linear features (i.e., the two features from of the data plus a single constant term). Plot the data points plus the linear separator.

(d) **(15-830 only)** Run a kernelized SVM on this task using a Gaussian kernel. You can pick the hyperparameters (bandwidth of kernel and regularization parameter) any way you want, but be sure to report the parameters you use. You might also want to scale the two features so that they are of similar magnitudes (by default, mean power consumption has a much smaller range than the durations). Because it is difficult to graph the separation boundary for non-linear classifiers, you can just report average training 0-1 loss of the resulting classifier.

3. **Sustainability on the brain [25 pts]**
You have a hypothesis: the more people think about sustainability the less energy they tend to use. It's of course hard to quantify "thinking about sustainability," but we can get some sense of this using tools like Google Trends (http://trends.google.com), which show how much a certain term is searched for.[1] Here is the Google Trends line for the search term "sustainability" (the markers are for news stories, which you can look at if you look at the webpage).

Clearly there seems to be some pattern to how much people search for sustainability (with the biggest drop coming right at Christmas, interestingly enough, but this happens for a lot of search terms). So can we use this to help us predict energy consumption? If so (and if indeed the direction favors increased search meaning less

[1]This tool created some waves in the medical community when researchers found that the search trends for symptoms of the flu were as predictive or more predictive of flu outbreaks than most existing models being used up to that point.

energy), then maybe we could claim that thinking about sustainability does cause us to use less energy. In this problem you'll go through the evaluation of whether or not this claim might be plausible.

You will use the data file `q3_data.txt` for this problem. Each line of this file contains average values of the Google trend, temperature, and demand for one week:

```
<Google trend> <average temp> <average demand>
```

(a) Plot the trend data versus average demand for all the data points.

(b) Using the trend data alone (not the temperature data), use linear regression to predict electrical demand given the trend data. You can use just linear regression here (adding a constant-valued feature to the input). Evaluate the performance of this model using 70% of the data as the training set and 30% of the data as a test set. Report mean squared error along with 95% confidence intervals on the test set. Also report mean squared error and 95% confidence intervals for a baseline prediction that just predicts the mean demand from the test set (this is essentially linear regression with *just* the constant term).

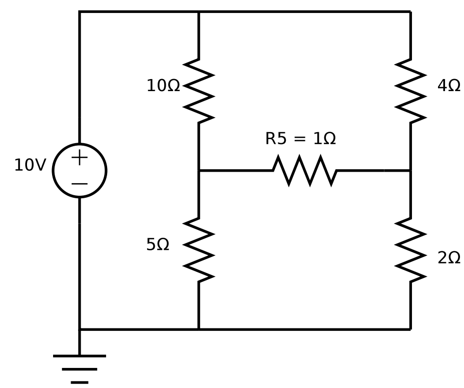
(c) Use both the trend and temperature data to predict demand, and compare this to using just temperature to predict demand, using the same evaluation methodology as above (i.e., for each method, report mean squared error and 95% confidence intervals on a test set). For the temperature, you can use quadratic features (temperature and temperature squared).

(d) What can you reasonably conclude from this study? Note that for the purposes of this problem you can conclude that predictor A is "significantly" better than predictor B if the mean squared error of predictor B lies above the 95% confidence interval of predictor A (i.e., it is enough for the mean of predictor B to lie outside A's error bars; the error bars do not have to be completely separate).

4. **Simple Circuits in Qucs [20 pts]**

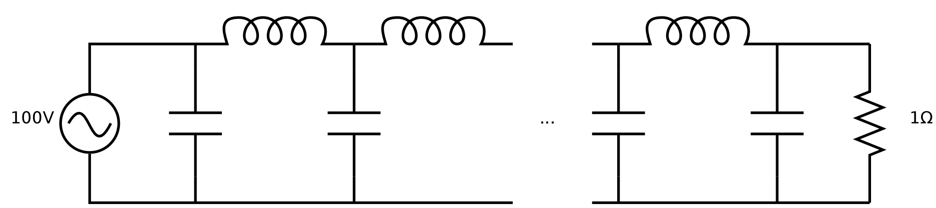
(a) Download and install Qucs, available at: `http://qucs.sourceforge.net/`. There are pre-built packages available for OS X and Windows. Ubuntu also has a package you can download directly `apt-get install qucs`. For all the questions in this section, you should submit a printout of the Qucs sheet (including the table with the answer).

(b) **Wheatstone bridge**. For the following circuit, determine the amount of current flowing through the resistor R_5 (even if you know how to do this analytically, for this question you'll need to solve it by laying out the circuit in Qucs).



The steps for doing this are as follows: 1) lay out the circuit with the power source, resistive elements, and ground connection; 2) place a current problem in line with the R_5 resistor; 3) place the DC simulation block, and click the simulate button (the gear icon on the top of the screen); 4) place a tabular diagram, then double click the current probe in the list to add its reading to the table.

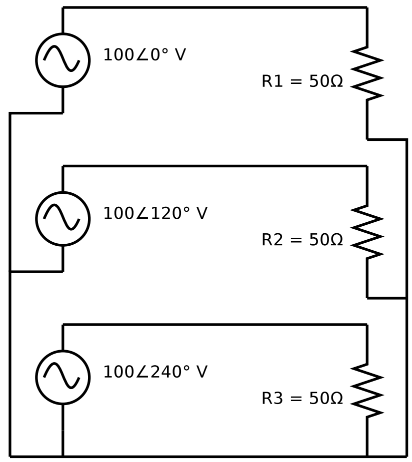
(c) **Transmission lines**. A simple model of a transmission line, that is slightly more accurate than the purely inductive version we have discussed in class, is the "pi" model, which consists of inductors in series and capacitors in parallel:



What is the (complex) voltage across the resistor for 1, 3, and 5 inductors. You can assume all inductors and capacitors have inductance and capacitance of 1 mH and 1 mF respectively (you only need to include the Qucs printout for one of these configurations).

For the AC circuit, you'll want to use an AC simulation block, then set it at a constant frequency of 60Hz (and be sure to change the AC voltage source frequency to 60Hz as well).

(d) **Three-phase power with unbalanced load** Consider a three-phase power system in a wye configuration (with explicit return wire) but with a potentially unbalanced load:



Confirm for the above setup that the current following flowing through the return line is zero amperes. What is the current flowing through the return line when $R_3 = 40\Omega$? Does the current flowing through R_1 change when we change R_3?.

5. **Three-phase power [15 pts]**

   (a) The model of a transformer we showed in the slides is a single-phase system, requiring a return line for the transformer. Show how you could build a three-phase transformer that could equally change the voltage of all three phases, without the need for any return lines. You should draw construct this using a combination of ordinary (single phase) transformers. (Hint: arrange a transformer like a delta or wye connection).

   (b) **(15-830 only)** We showed in class that three-phase power is the smallest number of phases for which (1) voltages/currents sum to zero (assuming three balanced loads and all currents in phase with voltage) and (2) power sum to a constant amount throughout the cycle. Is this also true of four phase power? For which polyphase (more than one phase) systems does this hold? To do this part rigorously, you'll have to make use of several trigonometric identities like the double angle formula; you can use any such identity you can find, but be sure to cite the source where you found it.