Assignment for "Supervised Learning" - Curtis Baker

Task 1. Single-layer classifier - Using provided code and data ("assign_classifier_gradDesc"):

A. Modify the code to graph the learning curve - loss function vs. the iteration number. Experiment with different values of learning rate. Show 3 examples of learning curves, when the learning rate is: too small; a value that works well; and too large. Describe in words, what in general are the consequences of it being too small or too large ? For this example, what is the largest value of learning rate that works well ?

B. Now modify the code to also graph all three weights vs. the iteration number, and illustrate what happens with a good value of the learning rate. Comment on what happens to the weight values during learning - what can you see in these curves, that might indicate overfitting ?

Task 2. Single-layer classifier with regularization

A. Using the code and data from Task 1, modify the updating of the weights, consistent with a penalty on the sum-squared weights. Take care to do this correctly - see MacKay, section 39.4, Figure 39.5.

B. Experiment with different values of the hyperparameter (alpha) - I suggest you try values of alpha between 0.0 and 1.0. Beware that for larger values of alpha, you will need smaller values of the learning rate in order to get sensible behaviour (also values between 0 and 1 might be a good range to explore). What is the effect of increasing the alpha value on the weight values, and on the overall results ? Show relevant Figures to illustrate what happens for an alpha value that is effective against overfitting.

Task 3. Receptive field (RF) estimation using regression with early stopping - use provided code to do this for a simulated model of a 1-d receptive field profile ("*assign_1d_RF_sysIdent_overfit.m*").

A. Modify the code to divide up the data into two sets, for Training (70%) and Validation (30%). Use the Training set for the iterative loop that learns the best estimate of the RF. Plot the learning curves, i.e. error (loss) vs. iterations, for both Training and Validation datasets (superimposed, with different line types). On this plot, indicate the best place to stop, for "early stopping". Show the full learning curve, well beyond the optimal place to stop.

B. Now at this optimal place to stop, plot the “learned” (estimated) receptive field, and the actual (model) receptive field profile, superimposed using different line and/or symbol types.

Be sure to indicate in the report, the learning rate and number of iterations that you used, and the resulting error (loss).

Task 4. RF estimation, comparing regression vs. correlation for different stimuli - use provided code ("assign_2d_RF_sysIdent") which uses a scaled conjugate gradient ridge-regression algorithm ('scg', from the provided *netlab* toolbox) that automatically optimizes the learning rate. The program provides different options for two types of visual stimuli and two types of analysis algorithm, for estimation of a 2d spatial receptive field.

A. The provided code partitions the data into two sets, for training and validation. Modify the code to provide a third "Test" partition, for final evaluation of the trained model performance, using VAF (variance accounted for). Be sure to use each of these datasets at the correct places in your code.

B. For scg-regression and white noise stimuli, modify the code to systematically search for the best alpha value to use for regularization. Graph VAF against alpha for the training dataset, on a semilogx plot (I suggest testing log-spaced values of alpha between 0.1 and 10000). Do this also for natural image stimuli. Show a figure with 4 subplots: the actual simulated model RF, and the estimated RF for three values of alpha (too small, optimal, much too large) - plot them all on the same z-scale, so they can be compared. Discuss the effects of the penalty value (alpha) being too small or too large.

C. Now evaluate the results using cross-correlation instead of regression - in this case, there is no ridge regression and therefore no alpha value - just run it and compare the results (estimated RF and VAF), for both white noise and natural image stimuli. Discuss relative advantages and disadvantages of regression vs. correlation methods for system identification of neuronal receptive fields.

Assignment guidelines:

The above exercises should be done in standard Matlab. You must provide a written report, as a PDF file. The report should be between 6 and 8 pages (including Figures), and easy to read (font size=12, margins = 1”). There should be a text "narrative", that briefly tells what you did for each part, what the results were, and your interpretation. Illustrate key results such as graphs or plots with Figures that are embedded within the main text narrative. Label axes of each plot and give each Figure a short “legend”, describing what it shows and indicating which of your Matlab script produced it, as well as any key parameter values. For each part of the Assignment, describe in the main text what you found, and discuss why (i.e., "what it means") - *it is not sufficient to simply show Matlab Figures, with only legends or with little or no explanatory text*. The report should not only describe what you find, but very importantly, should discuss what you think the results mean or illustrate.

Also please provide the Matlab code in a form that I can run - it should be executable scripts (not functions) that reproduce the graphs or plots for each Task in your report, so I can see what you have done - with a separate script for each Task. Please do not forget to include any custom or non-standard functions. The Matlab code must be provided *in addition to* the above written report; comments in the M-files are very welcome, but they do not substitute for the report. Be sure to tell the version of Matlab you are using, since some functions are supported only in some versions not others.

All of this should be packaged in a single archive file (.zip or .hqx). Please be sure to include your last name, as the prefix to the file name - e.g. Smith_603_Assign.zip