1 Textbook exercises

The textbook has many programming exercises suitable for reviewing for Test 2:

Chapter 5 17, 18, 19, 22, 24, 26, 28, 29, 30, 32, 33, 36, 40, 45

Chapter 12 1, 3, 7, 8 (try to generalize your function so that it can compute a moving average of every n elements), 13, 19, 25

2 Sample written questions

1. Explain what the following MATLAB statements do assuming `x` is a vector of numbers:
   (a) `x > 0`
   (b) `x > 0 & x < 10`
   (c) `sum(isinf(x) | isnan(x))`
   (d) `x(x < 0)`
   (e) `x(isnan(x)) = []`

   —

2. The textbook suggests that you should always preallocate a vector or matrix if you can, rather than append elements or rows or columns to grow the vector or matrix. Why is this a good idea?
3. Why does the solution for the Collatz problem from Lab 5 not preallocate a vector for the sequence?
4. Consider the following MATLAB `for` loop:

```
for k = 1:n
  % something happens here
end
```

   Rewrite the loop using a `while` loop.

5. Consider the following MATLAB `while` loop:

```
done = 0;
k = 1;
while ~done
  % something happens here
  if k > n
    done = 1;
  end
end
```

   Rewrite the loop using a `for` loop.

6. Suppose you have a vector `x` of length `n` unique numbers (that are not NaN, Inf, or -Inf); how many elements of `x` are:
   (a) less than the first quartile of `x`?
   (b) less than the second quartile of `x`?
   (c) greater than the third quartile of `x`?
7. The function `trimmean` was discussed in the lecture slides; why does MATLAB not have a `trimmedian` function?
8. Your transcript lists the average grade for each course that you have completed; I would argue that the average grade is often not a useful summary statistic for university courses. Give an example of when the mean grade for a course provides misleading information to the casual reader.

3 More sample programming questions

1. Look at the documentation for `boxplot` and determine for what values does boxplot consider a point to be an outlier (look for the `'whisker'` parameter). Write a function that returns the indices of all of the outliers in a vector of measurements using the same criteria that `boxplot` uses. For example, the vector `x = [47 47 2 12 59 85 47 60 36 93 51 55]` has outliers at indices `[3 4 10]`

2. Write a function that computes the numbered grid for a game of minesweeper. See the Overview section of the Wikipedia page `http://en.wikipedia.org/wiki/Minesweeper_%28video_game%29` for the values in the grid. Start with the following function:

```
function B = minesweeper(A)
%MINESWEEPER Computes the numbered grid for a game of minesweeper
%   B = MINESWEEPER(A) computes the numbered grid for a game
%   of minesweeper. A is grid containing zeros (no mines) and -1
%   (mines). Each element in B contains the number of adjacent
%   grid positions containing a mine in A, or -1 if there is a mine
%   at that position in A.
%
%   For simplicity, the first and last rows and columns of B do
%   not need to be computed (e.g, they can be all zeros).

[nrows, ncols] = size(A);
B = A;
for row = 2:(nrows - 1)
    for col = 2:(ncols - 1)




    end
end
```

3. (Too hard for a test; also requires material not included in Test 2) One problem with using the method of least squares to find a best fit line to a set of data points is that the least squares estimate is very sensitive to erroneous data points. An alternative to least squares is to compute the line that minimizes the median of the squared residual errors; unfortunately, the median cannot be differentiated so calculus techniques cannot be used to find such a line.

   One approach to finding the least median of squares line is to realize that only two points are needed to define a line in two dimensions. A possible algorithm for computing the best fit line examines every combination of two unique data points and:

   - define the line L that passes through the two data points
   - compute the residual errors for all of the data points using L
   - compute the median of the squared residual errors

   The best fit line is the one that has the smallest median of the squared residual errors.

   Write a function that computes the least median of squares best fit line. Fitting a line to two points and computing the residual errors can all be done using `polyfit` and `polyval`; the interesting part of the question is computing all of the combinations of two data points.