Unconstrained Optimization

Landon Wright

February 3, 2018

1 Program Description

2 Testing Results

2.1 Function 1 Results

$Starting_x$	$Starting_y$	$Starting_z$	Function Val	$SearchDirection_x$	$SearchDirection_y$	$SearchDirection_z$	Step
10.000000	10.000000	10.000000	154.343825	-0.994269	0.064146	-0.085528	7.818
2.226306	10.501529	9.331295	38.883122	0.033560	-0.572300	-0.819357	12.52
2.646680	3.332821	-0.932087	1.547244	-0.976543	0.155690	-0.148744	2.512
0.193426	3.723944	-1.305758	-12.992294	0.123689	-0.159847	-0.979362	4.380
0.735284	3.023680	-5.596172	-18.666244	-0.981909	0.122884	-0.144067	0.979
-0.226799	3.144083	-5.737330	-20.913770	0.104303	-0.283997	-0.953135	1.721
-0.047292	2.655322	-7.377682	-21.804129	-0.980760	0.129642	-0.145955	0.388
-0.427967	2.705642	-7.434333	-22.157079	0.108735	-0.258162	-0.959963	0.681
-0.353812	2.529581	-8.089008	-22.296993	-0.980997	0.128234	-0.145603	0.153
-0.504753	2.549312	-8.111412	-22.352459	0.107833	-0.263520	-0.958608	0.270
-0.475600	2.478069	-8.370572	-22.374447	-0.980948	0.128524	-0.145677	0.060
-0.535434	2.485908	-8.379457	-22.383164	0.108020	-0.262414	-0.958890	0.107
-0.523857	2.457784	-8.482226	-22.386619	-0.980958	0.128464	-0.145662	0.024
-0.547577	2.460891	-8.485748	-22.387989	0.107981	-0.262642	-0.958832	0.042
-0.542990	2.449732	-8.526485	-22.388532	-0.980956	0.128477	-0.145665	0.009
-0.552393	2.450963	-8.527882	-22.388747	0.107989	-0.262595	-0.958844	0.016
-0.550574	2.446541	-8.544031	-22.388833	-0.980957	0.128474	-0.145665	0.003
-0.554302	2.447029	-8.544585	-22.388867	0.107988	-0.262605	-0.958842	0.006
-0.553581	2.445275	-8.550987	-22.388880	-0.980957	0.128475	-0.145665	0.001
-0.555059	2.445469	-8.551206	-22.388885	0.107988	-0.262603	-0.958842	0.002
-0.554773	2.444774	-8.553744	-22.388888	-0.980957	0.128475	-0.145665	0.000
-0.555359	2.444851	-8.553831	-22.388888	0.107988	-0.262603	-0.958842	0.001
-0.555245	2.444575	-8.554838	-22.388889	-0.980957	0.128475	-0.145665	0.000
-0.555477	2.444605	-8.554872	-22.388889	0.107988	-0.262603	-0.958842	0.000
-0.555433	2.444496	-8.555271	-22.388889	-0.980957	0.128475	-0.145664	0.000
-0.555525	2.444508	-8.555285	-22.388889	0.107988	-0.262603	-0.958842	0.000
-0.555507	2.444465	-8.555443	-22.388889	-0.980957	0.128475	-0.145665	0.000
-0.555543	2.444470	-8.555448	-22.388889	0.107978	-0.262602	-0.958844	0.000
-0.555536	2.444453	-8.555511	-22.388889	-0.980958	0.128478	-0.145654	0.000
-0.555551	2.444454	-8.555513	-22.388889	0.107983	-0.262603	-0.958843	0.000
-0.555548	2.444448	-8.555538	-22.388889	-0.980958	0.128478	-0.145653	0.000
-0.555554	2.444448	-8.555539	-22.388889	0.108024	-0.262608	-0.958837	0.000
-0.555553	2.444446	-8.555549	-22.388889	-0.980962	0.128488	-0.145616	0.000

$Starting_x$	$Starting_y$	$Starting_z$	Function Val	$SearchDirection_x$	$SearchDirection_y$	$SearchDirection_z$	St
10.0000	10.0000	10.0000	154.3438	-0.9943	0.0641	-0.0855	7.
2.2263	10.5015	9.3313	-14.3999	-0.1593	-0.5491	-0.8204	18
-0.7467	0.2560	-5.9770	-22.3889	0.0564	0.6460	-0.7612	3.

$Starting_x$	$Starting_y$	$Starting_z$	Function Val	$SearchDirection_x$	$SearchDirection_y$	$SearchDirection_z$	St
10.0000	10.0000	10.0000	154.3438	-0.9943	0.0641	-0.0855	7.
2.2263	10.5015	9.3313	-14.3999	-0.1593	-0.5491	-0.8204	18
-0.7467	0.2560	-5.9770	-22.3889	0.0564	0.6460	-0.7612	3.

2.2 Rosenbrock Function Results

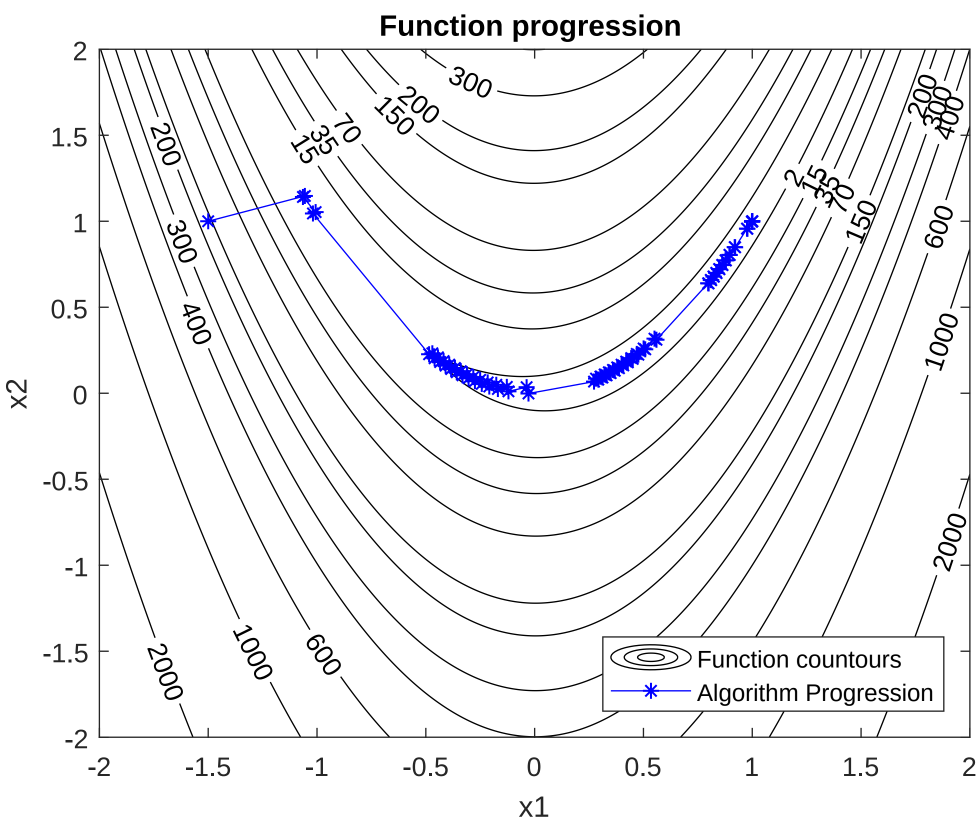


Figure 1: Progression of steepest descent algorithm

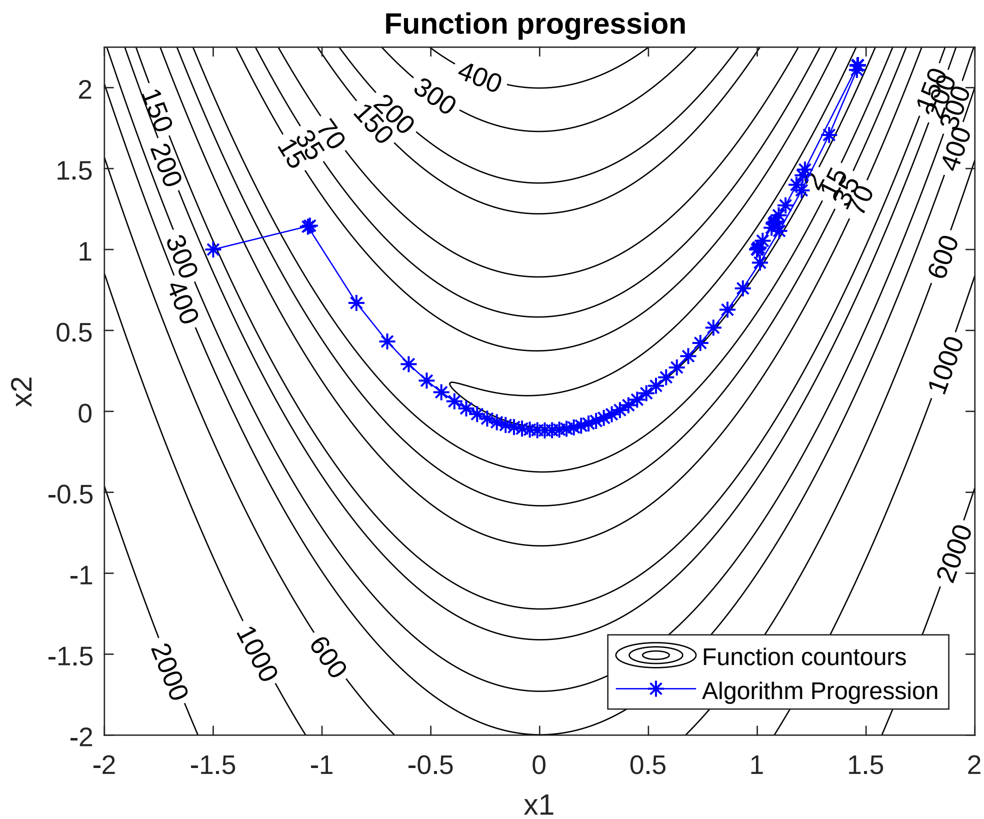


Figure 2: Progression of conjugate gradient algorithm

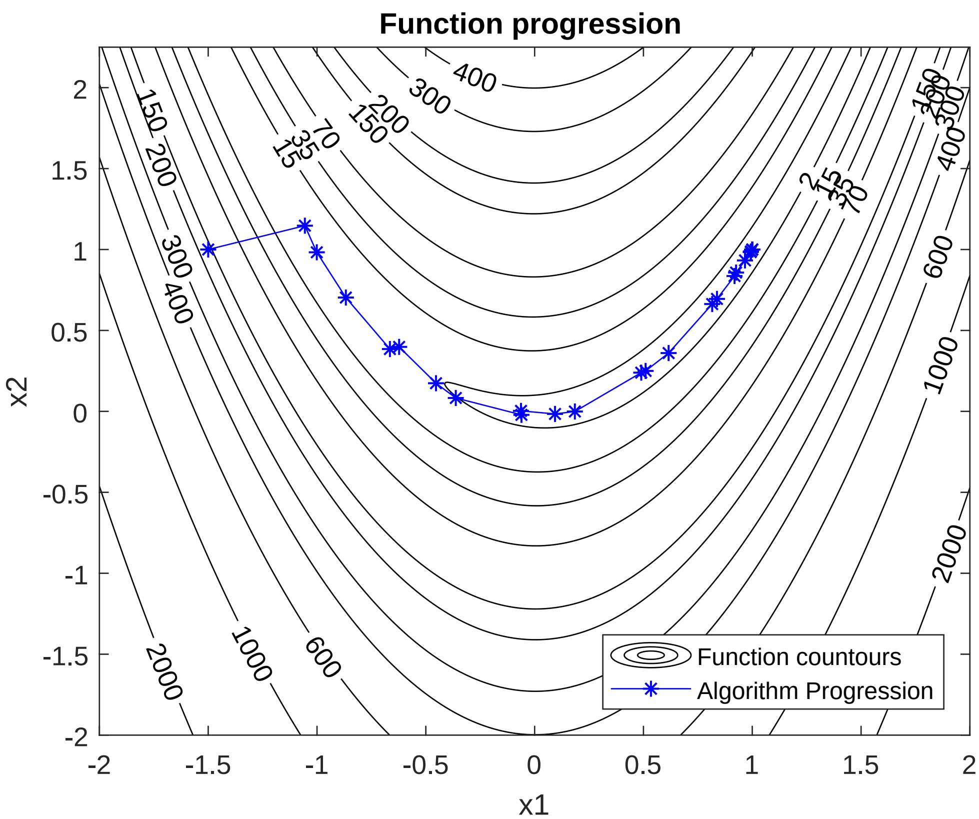


Figure 3: Progression of quasi-Newton algorithm

3 Matlab Code

3.1 Fminun Routine

```
function [xopt, fopt, exitflag] = fminun(obj, gradobj, x0, stoptol, algoflag)

  % get function and gradient at starting point
  global nobj;
  [n,~] = size(x0); % get number of variables
  f = obj(x0); % get the value of the function at x0
  grad = gradobj(x0);
  x = x0;
  fOld = inf;
  xOld = zeros(n, 1);
  %set starting step length
  alpha = 0.0005;
  incrementCounter = 0;
  gradOld = ones(n,1);
  sOld = zeros(n,1);
  N = eye(n);
  saveMat = table;


  while (any(abs(grad(:)) > stoptol))
    incrementCounter = incrementCounter + 1
    if (nobj > 500)
      xopt = nan;
      fopt = nan;
      exitflag = 1
      return
    end

    if (algoflag == 1)     % steepest descent
      s = srchsd(grad);
      % find the proper alpha level
      % function [alphaPrime] = aPrime(obj, gradobj, s, f, x)
      alphaPrime = aPrime(obj, gradobj, s, f, x);

    end

    if (algoflag == 2)
      % use conjugate gradient method
      % check that it's not the first round / we wont divide by 0
      if (gradOld' * gradOld == 0)
        sMessy = -grad;
      else
        % calculate the beta term
        betaCorrection = (grad' * grad) / (gradOld' * gradOld);
        sMessy = -grad + betaCorrection * sOld;
        sOld = sMessy;
        % normalize the s vector
        s = sMessy / norm(sMessy);
        alphaPrime = aPrime(obj, gradobj, s, f, x);
      end
```

```

    end

    if (algoflag == 3)
      % use quasi-Newton method
      gammaK = grad - gradOld;
      deltaX = x - xOld;
      if (incrementCounter > 1 & deltaX' * gammaK > 0)
        t1 = 1 + ((gammaK' * N * gammaK) / (deltaX' * gammaK));
        t2 = (deltaX * deltaX') / (deltaX' * gammaK);
        t3 = (deltaX * gammaK' * N + N * gammaK * deltaX') / (deltaX' * gammaK);
        N = N + t1 * t2 - t3
      end
      s = -N * grad;
      s = s / norm(s);
      alphaPrime = aPrime(obj, gradobj, s, f, x);

    end


    % take a step
    xnew = x + alphaPrime*s;
    fnew = obj(xnew);
    gradOld = grad;
    grad = gradobj(xnew);
    fOld = f;
    xOld = x;
    f = fnew;
    x = xnew;
    newRow = {xOld', fnew, s', alphaPrime, nobj};
    saveMat = [saveMat; newRow];


  end
  grad
  xopt = xnew;
  fopt = fnew;
  exitflag = 0;
  % saveMat.Properties.VariableNames = {'Starting_Point', 'Function_Value', ...
  %     'Search_Direction', 'Step_Length', 'Number_of_Objective_Evaluations'};
  toSave = table2array(saveMat);
  fout = fopen(sprintf('output%d.csv', algoflag),'w');
  fprintf(fout, '%s, %s, %s, %s, %s, %s, %s, %s, %s\r\n'...
  , '$Starting_x$', '$Starting_y$', '$Starting_z$', 'Function Val', ...
   '$Search Direction_x$', '$Search Direction_y$', '$Search Direction_z$', ...
   'Step Length', 'Number of Objective Evaluations');
  fprintf(fout, '%8.6f, %8.6f, %8.6f, %8.6f, %8.6f, %8.6f, %8.6f, %8.6f, %8.6f \r\n', toSave');
  fclose(fout);
  % writetable(saveMat, sprintf('output%d.csv', algoflag), 'precision', '%10.4f');
  % write the matrix to a file

end

% get steepest descent search direction as a column vector
```

```
function [s] = srchsd(grad)
  mag = sqrt(grad'*grad);
  s = -grad/mag;
end
```

3.2 Alpha* line search

```
function [alphaPrime] = aPrime(obj, gradobj, s, f, x)
  minVal = f;
  lastStepVal = f;
  alphas = [0, f];
  testStep = 0.1;
  xTest = x;
  incrementer = 2;
  iterTestStep = testStep;
  while (minVal >= lastStepVal)
    % calculate at guessed testStep
    xTest = x + iterTestStep * s;
    fTest = obj(xTest);
    if(fTest >= f & incrementer < 3)
      % disp("Step size to big, recalculating");
      minVal = f;
      lastStepVal = f;
      iterTestStep = iterTestStep / 10;
      alphas = [0, f];
      incrementer = 2;
      continue;
    end
    % add it to the stored list
    alphas(incrementer,1) = iterTestStep;
    alphas(incrementer,2) = fTest;
    % increment things for the next loop
    minVal = min(minVal, fTest);
    lastStepVal = fTest;
    incrementer = incrementer + 1;
    alphaOpt = iterTestStep;
    iterTestStep = iterTestStep * 2;
  end
  % take the half step between the last two steps
  iterTestStep = (alphas(end, 1) + alphas(end-1, 1)) / 2;
  xTest = x + iterTestStep * s;
  fTest = obj(xTest);
  % store the values of the intermediate step
  alphas(end + 1, :) = alphas(end, :);
  alphas(end - 1, :) = [iterTestStep, fTest];
  % find the index of the minimum function value
  [minVal, minIdx] = min(alphas(:,2));
  % get the three alpha and function values
  alpha2 = alphas(minIdx, 1);
  f2 = alphas(minIdx, 2);
  alpha1 = alphas(minIdx - 1, 1);
  f1 = alphas(minIdx - 1, 2);
  alpha3 = alphas(minIdx + 1, 1);
  f3 = alphas(minIdx + 1, 2);
```

```
    [alpha1, alpha2, alpha3];
    % calculate the optimum alpha value
    deltaAlpha = alpha2 - alpha1;
    alphaPrime = (f1 * (alpha2^2 - alpha3^2) + f2 * (alpha3^2 - alpha1^2) ...
          + f3 * (alpha1^2 - alpha2^2)) / (2 * (f1 * ...
          (alpha2 - alpha3) + f2 * (alpha3 - alpha1) + ...
          f3 * (alpha1 - alpha2)));
end
```

3.3 Driver

```
function [] = fminunDriv()
%----------------Example Driver program for fminun------------------
    clear;

    global nobj ngrad
    nobj = 0; % counter for objective evaluations
    ngrad = 0.; % counter for gradient evaluations
    x = [1.; 1.]; % starting point, set to be column vector
    x1 = [10; 10; 10]; % starting point for function 1
    xRosen = [-1.5; 1];
    algoflag = 1; % 1=steepest descent; 2=conjugate gradient; 3=BFGS quasi-Newton
    stoptol = 1.e-5; % stopping tolerance, all gradient elements must be < stoptol


    % ---------- call fminun----------------
    [xopt, fopt, exitflag] = fminun(@obj1, @gradobj1, x1, stoptol, algoflag);

    xopt
    fopt

    nobj
    ngrad
end

 % function to be minimized
 function [f] = obj(x)
    global nobj
    %example function
    %min of 9.21739 as [-0.673913, 0.304348]T
    f = 12 + 6.*x(1) - 5.*x(2) + 4.*x(1).^2 -2.*x(1).*x(2) + 6.*x(2).^2;
    nobj = nobj +1;
 end

% get gradient as a column vector
 function [grad] = gradobj(x)
    global ngrad
    %gradient for function above
    grad(1,1) = 6 + 8.*x(1) - 2.*x(2);
    grad(2,1) = -5 - 2.*x(1) + 12.*x(2);
    ngrad = ngrad + 1;
 end

 % function 1 to be optimized on the homework
```

```
function [f] = obj1(x)
  global nobj
  f = 20 + 3 .* x(1) - 6 .* x(2) + 8 .* x(3) + 6 .* x(1).^2 - 2 .* x(1) .* x(2) ...
       - x(1) .* x(3) + x(2).^2 + 0.5 .* x(3).^2;
  nobj = nobj + 1;
end

% gradient of function 1
function [grad] = gradobj1(x)
  global ngrad
  grad(1,1) = 3 + 12 .* x(1) - 2 .* x(2) - x(3);
  grad(2,1) = -6 - 2 .* x(1) + 2 .* x(2);
  grad(3,1) = 8 - x(1) + x(3);
  ngrad = ngrad + 1;
end


% function 1 to be optimized on the homework
function [f] = objRosen(x)
  global nobj
  f = 100 .* (x(2) - x(1).^2).^2 + (1-x(1)).^2;
  nobj = nobj + 1;
end

% gradient of function 1
function [grad] = gradobjRosen(x)
  global ngrad
  grad(1,1) = 2 .* (200 .* x(1).^3 - 200 .* x(1) .* x(2) + x(1) - 1);
  grad(2,1) = 200 .* (x(2) - x(1).^2);
  ngrad = ngrad + 1;
end
```

ME 575
Homework #3 Unconstrained Optimization
Due Feb 7 at 2:50 p.m.

Description
Write an optimization routine in MATLAB (required) that performs unconstrained optimization. Your routine should include the ability to optimize using the methods,

- steepest descent
- conjugate gradient
- BFGS quasi-Newton

Your program should be able to work on both quadratic and non-quadratic functions of *n* variables. To determine how far to step, you may use a line search method (such as the quadratic fit given in the notes), a trust region method, or a combination of these. You will be evaluated both on how theoretically sound your program is and how well it performs.

You should use good programming style, such as selecting appropriate variable names, making the program somewhat modular (employing function routines appropriately) and documenting your code.

You will provide test results for your program on the two functions given here. In addition, you will turn in your function so we can test it on other functions or on different starting points.

Testing
1) Test your program on the following quadratic function of three variables:

$$f = 20 + 3x_1 - 6x_2 + 8x_3 + 6x_1^2 - 2x_1x_2 - x_1x_3 + x_2^2 + 0.5x_3^2$$

The conjugate gradient and quasi-Newton methods should be able to solve this problem in three iterations. Show your data for each iteration (starting point, function value, search direction, step length, number of evaluations of objective) for these two methods starting from the point $\mathbf{x}^T = [10, 10, 10]$. In addition, give data for five steps of steepest descent. At the optimum the absolute value of all elements of the gradient vector should be below 1.e-5. Indicate the total number of objective evaluations and gradient evaluations taken by each method.

2) Test your program on the following non-quadratic function (Rosenbrock's function) of two variables:

$$f = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

Starting from the point $\mathbf{x}^T = [-1.5, 1]$, show the steps of the methods on a contour plot. Show 20 steps of steepest descent. Continue with the conjugate gradient and quasi-Newton methods until the optimum is reached at $\mathbf{x}^T = [1, 1]$ and the absolute value of all elements of the gradient vector is below 1.e-3. Indicate the total number of objective evaluations and gradient evaluations taken by each method.

MATLAB Stuff

Your function will be called `fminun`. It will receive and pass back the following arguments:

```
[xopt, fopt, exitflag] = fminun(@obj, @gradobj, x0, stoptol, algoflag);

Inputs:
@obj = function handle for the objective we are minimizing (we will use
obj) The calling statement for obj looks like,

f = obj(x);

@gradobj = function handle for the function that evaluates gradients of
the objective (we will use gradobj, see example) The calling statement for
gradobj looks like,

grad = gradobj(x);

Note that grad is passed back as a column vector.

x0 = starting point (column vector)

stoptol = stopping tolerance. I will set this to 1.e-3, unless this proves
too restrictive. The absolute value of all elements of your gradient
vector should be less than this value at the optimum.

algoflag = 1 for steepest descent, =2 for conjugate gradient, =3 for
quasi-Newton.

Outputs:
xopt = optimal value of x (column vector)

fopt = optimal value of the objective

exitflag = 0 if algorithm terminated successfully; otherwise =1. Your
algorithm should exit (=1) if it has exceeded more than 500 evaluations of
the objective.
```

Attached is an example “driver” routine and example `fminun` routine to get you started. You can copy these from Learning Suite > Content > MATLAB Examples.

Grading

Grading: Your routine will be graded based on two criteria: 1) Soundness of your methodology and implementation as evidenced by your write-up and code (50%), and performance on test functions (50%). The performance score will be based 70% on accuracy (identifying the optimum) and 30% on efficiency (number of objective evaluations plus *n**number of gradient evaluations).

Turn in, in one report through Learning Suite:

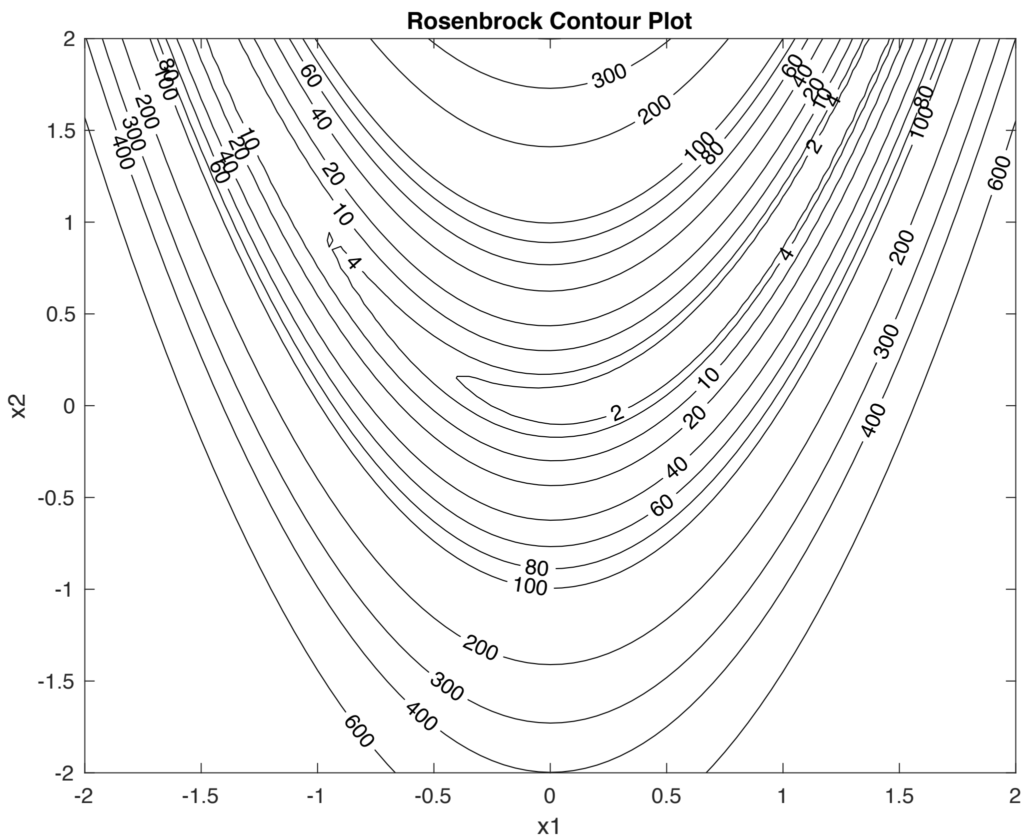
1. A brief written description of your program, including discussion of each of the three methods. This should not be longer than a page (single-spaced). You may include equations if you wish. Discuss how you implemented the methods.
2. The requested results from testing on the two functions.

3. Hardcopy of your MATLAB code.

In addition, email your MATLAB code to Jacob Greenwood (jacobgwood@gmail.com) . Additional information about this will be provided.

Suggestions:

Get started now! Start with a relatively simple routine and work forward, always having a working program. I would suggest you start off with a relatively straight-forward step length approach (such as the quadratic fit given in class) and then you can get more fancy after you have a program working for all three methods, if you want to. A working, straightforward program is much better than a non-working, fancy program.



Rosenbrock's function. The optimum is at $(\mathbf{x}^*)^T = [1, 1]$ where $f^* = 0$. Start at the point, $(\mathbf{x}^0)^T = [-1.5, 1]$

Driver Routine:

```
%----------------Example Driver program for fminun------------------
    clear;

    global nobj ngrad
    nobj = 0; % counter for objective evaluations
    ngrad = 0.; % counter for gradient evaluations
    x0 = [1.; 1.]; % starting point, set to be column vector
    algoflag = 1; % 1=steepest descent; 2=conjugate gradient; 3=BFGS quasi-Newton
    stoptol = 1.e-3; % stopping tolerance, all gradient elements must be < stoptol


    % ---------- call fminun----------------
    [xopt, fopt, exitflag] = fminun(@obj, @gradobj, x0, stoptol, algoflag);

    xopt
    fopt

     nobj
    ngrad

     % function to be minimized
     function [f] = obj(x)
        global nobj
        %example function
        f = 12 + 6*x(1) - 5*x(2) + 4*x(1)^2 -2*x(1)*x(2) + 6*x(2)^2;
        nobj = nobj +  1;
     end

    % get gradient as a column vector
     function [grad] = gradobj(x)
        global ngrad
        %gradient for function above
        grad(1,1) = 6 + 8*x(1) - 2*x(2);
        grad(2,1) = -5 - 2*x(1) + 12*x(2);
        ngrad = ngrad + 1;
     end
```

Example fminun function:

```
function [xopt, fopt, exitflag] = fminun(obj, gradobj, x0, stoptol, algoflag)

   % get function and gradient at starting point
   [n,~] = size(x0); % get number of variables
   f = obj(x0);
   grad = gradobj(x0);
   x = x0;

   %set starting step length
   alpha = 0.5;

   if (algoflag == 1)     % steepest descent
      s = srchsd(grad)
   end

   % take a step
   xnew = x + alpha*s;
   fnew = obj(xnew);

   xopt = xnew;
   fopt = fnew;
   exitflag = 0;
end

% get steepest descent search direction as a column vector
function [s] = srchsd(grad)
   mag = sqrt(grad'*grad);
   s = -grad/mag;
end
```