Question 1: Consider the problem of minimizing the function:

$$f(x_1, x_2) = 5x_1^2 + 4x_1x_2 + x_2^2 - 5x_1 - 3x_2 + 6$$

a) Convert the function to the quadratic form:

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} - \mathbf{x}^T\mathbf{b} + c$$

where $\mathbf{x} = \{x_1, x_2\}^T$ and $\mathbf{Q}$ is a symmetric positive definite matrix.

b) Find analytically the exact solution to the problem.

c) Solve the problem using the steepest descent approach starting from the initial guess $\mathbf{x}^0 = \{0,0\}^T$. Use 15 iterations.

d) Solve the problem using the conjugate gradient approach starting from $\mathbf{x}^0 = \{0,0\}^T$.

e) In a plot on the $x_1 - x_2$ plane, compare the convergences of $\mathbf{x}$ with iterations obtained using the steepest descent and conjugate gradient approaches. In the same plot also show the exact solution obtained analytically. Comment on your results.

Question 2: The MATLAB data file HW2_Prob2_data.mat contains 1000 measurements of three variables: independent variables x and y; and dependent variable z. While the measurements of the independent variables are exact, those of the dependent variable are noisy (have some measurement errors). It is desired to fit a model to the data, and the following form of the model is chosen:

$$z = c(x + \alpha)^a y^b$$

where a, b, c and α are parameters of the model. 500 sets of measurements will be used to identify the parameters of the model – this will be referred to as the training data set. The remaining 500 sets of measurements will be used to test the performance of the model, with the identified model parameters, in predicting the values of the dependent variable z – this will be referred to as the testing data set.

a) Use the training dataset (*x*_train, *y*_train, *z*_train) from the MATLAB data file to identify the parameters a, b, c and α using the Gauss-Newton approach discussed in class. Start from the initial guess: $a = 1$, $b = 1$, $c = 1$ and $\alpha = 1$. Perform 50 iterations of the Gauss-Newton approach. Show the convergence of the different parameters with iterations using a plot of the parameter values at every iteration. What are the converged values of the different parameters? How many iterations are required for convergence?

b) Find the parameters a, b, c and α, using the same training dataset and starting from the same initial values as in part (a), but now using the inbuilt MATLAB function *lsqnonlin*. How do the converged values of the different parameters compare against the values obtained in part (a)?

c) Use the parameter values identified in part (a) with the testing dataset (*x*_test, *y*_test) from the MATLAB data file to predict the values of z. How do the predicted values of z compare with the corresponding measured values of z (given as *z*_test in the MATLAB data file)? You can illustrate this comparison by a scatter plot with the measured values of z in the x-axis and corresponding predicted values of z in the y-axis.

d) For the predicted values of z in part (c), compute the relative prediction errors (in %), defined as:

$$e = \frac{\text{Predicted value of } z - \text{Corresponding measured value of } z}{\text{Corresponding measured value of } z} \times 100$$

Based on the 500 measurements in the testing data set, find the mean and standard deviation of the computed relative prediction errors. Do you think the predictions have any significant bias? Further, use *normplot* and/or *dfittool* in MATLAB to check if the prediction errors are normally distributed?

Suggestion: I think you can ease the computations involved in this problem by converting the original model: $z = c(x + \alpha)^a y^b$, to a model of reduced nonlinearity by taking log of both sides.