NUMERICAL OPTIMISATION
TUTORIAL 6/03/20
ASSIGNMENT 3 (submit by 11pm on Thursday 12/03)

Marta Betcke

EXERCISE 1

(a) Implement the BFGS method by modifying the `descentLineSearch` function. More help is provided inside MATLAB Grader.
Submit your solution via **MATLAB Grader**. **[10pt]**

(b) Make your implementation efficient as explained in the lecture i.e. avoid explicitly forming the inverse Hessian matrix H_k. Copy the code lines implementing the update of H_k into your report and briefly explain what makes the implementation efficient.
Submit your solution via **TurnitIn**. **[10pt]**

EXERCISE 2 [DEMO]

Implement the SR-1 method by modifying the `trustRegion` function. More help is provided inside MATLAB Grader. **Note:** Here you are not expected to provide an efficient implementation as it would require some changes to `solverCM2dSubspaceExt` which are out of scope at this point. **[0pt]**

EXERCISE 3

(a) Minimise the function
$$f(x, y) = (x - 3y)^2 + x^4$$

using BFGS (**Ex 1a**) and SR1 (**Ex 2**) methods starting from $x_0 = (10, 10)^T$. Evaluate the performance of the two methods. Visualise the paths traced by the iterates. Provide any parameter values used, if needed motivate your choices.

Provide convergence plots and compare theoretical vs practical convergence rates. Attempt to explain what you observe. *Hint: Pay attention to the conditioning of the Hessian and the approximations, step lengths and activity / inactivity of the trust region. Check the* `info` *tructure for relevant information.*
Submit your solution via **TurnitIn**. **[20pt]**

(b) Both implementations have an option (which is active) to return a sequence of matrices as a field of the `info` structure:

  (i) $\{H_k^{\mathrm{BFGS}}\}_{k \geq 0}$ when using BFGS,

  (ii) $\{B_k^{\mathrm{SR1}}\}_{k \geq 0}$ when using SR1.

Plot the error of these sequences obtained in **Ex 3a** with respect to the matrices they approximate. In particular, plot

  (i) $\{||I - H_k^{\mathrm{BFGS}} \nabla^2 f(x_k)||_2\}_{k \geq 0}$,

  (ii) $\{||I - (\nabla^2 f(x_k))^{-1} B_k^{\mathrm{SR1}}||_2\}_{k \geq 0}$,

and explain your results.
Submit your solution via **TurnitIn**. **[10pt]**

EXERCISE 4

Implement the Gauss-Newton method for solution of nonlinear least square problems. As Gauss-Newton is a line search method, it can be easiest implemented inside the function `descentLineSearch.m`. More help is provided in `Cody Coursework`.
Submit your implementation via MATLAB Grader. **[10pt]**

EXERCISE 5 [DEMO]

You are given an implementation of the Levenberg-Marquardt method. This implementation plugs in the Levenberg-Marquardt solver `solverCMlevenberg.m` into the trust region function `trustRegion.m`. Briefly answer the following questions about the solver in `solverCMlevenberg.m`

(i) What is the effect of the `for` loop in line 38 and what is the reason for sequential execution?

(ii) Explain the formula for calculating the L-M direction p.

(iii) Explain what is q and how it is calculated.

(iv) Explain what is λ and the formula used for its update.

[0pt]

EXERCISE 6

Consider a model
$$\varphi(x_1, x_2, x_3; t) = (x_1 + x_2 t^2) \exp(-x_3 t)$$
with parameters (x_1, x_2, x_3).

Simulate the measurements sampling this model for a fixed choice of parameters $(x_1, x_2, x_3) = (3, 150, 2)$ at 200 equi-spaced points in $t_i \in (0, 4]$ and adding Gaussian noise $n(t_i) \sim \mathcal{N}(0, \sigma^2)$ drawn from a normal distribution with 0 mean and standard deviation 5% of the maximal amplitude of the sampled model signal i.e. $\sigma = 0.05 \max_{t_i} |\varphi(t_i)|$,

$$\tilde{\varphi}(x_1, x_2, x_3; t_j) = \varphi(x_1, x_2, x_3; t_j) + n(t_j).$$

(a) Formulate the least-squares problem for fitting the model φ and derive its Jacobian.
Submit your solution via TurnitIn. **[15pt]**

(b) Estimate the parameters (x_1, x_2, x_3) from your simulated measurements using

   (i) Gauss-Newton method (implemented in **Ex 1**);

   (ii) Levenberg-Marquardt method (implementation provided on Moodle).

Evaluate the performance of both methods. Visualise the fit by plotting the estimated signal versus the measurements. Specify all relevant parameters, if needed motivate your choices.

Provide convergence plots and compare the theoretical and practical convergence rates. Attempt to explain what you observe. *Hint: Evaluate the factor* $\|[J^{\mathrm{T}} J(x^\star)]^{-1} H(x^\star)\|$.

Submit your solution via TurnitIn. **[25pt]**

Remark. The submission to *TurnitIn* should not exceed 8 pages. This is not a hard limit and there is no penalty for longer submissions. Avoid submitting code unless explicitly asked for and focus on explaining your results.