

IAM566 Computer Assignment I

Ismail Hakki Kocdemir, 2036051

November 2019

1

For the minimization of the following problem

$$f(x_1, x_2) = (1.5x_1 + x_1x_2)^2 + (2.5x_1 + x_1x_2^2)2 + (2.625x_1 + x_1x_2^3)^2,$$

Steepest descent with armijo condition, Newton's and BFGS methods are used for different parameter sets. The table below contains all experimental results for the initial point $x_0 = (3.1, 0.4)^T$:

	iterations			$\ x_k - x^*\ _2$			$\ \nabla f_k\ _2$		
	1e-2	1e-3	1e-4	1e-2	1e-3	1e-4	1e-2	1e-3	1e-4
Steepest Descent (Armijo)	4	7	108	0.0355	0.0356	0.0359	0.0074	3.1e-4	7.4e-5
Newton's Method	3	4	5	0.0400	0.0361	0.0361	0.0085	3.2e-4	4.1e-9
BFGS updated	3	5	6	0.0441	0.0361	0.0361	0.0085	3.2e-4	4.1e-9

We can see that, Newton's method and BFGS has taken less steps to reach the given tolerance in all cases compared to steepest descent, since they have quadratic convergence and steepest descent has linear convergence. We can also see slight increase in iterations in BFGS compared to Newton's. This is expected as BFGS approximates the hessian. We can also see that smallest stopping tolerance values in BFGS and Newton results in closer points to actual minimum, which is not the case in steepest descent due to the step size regulation and the fact that linear convergence leads to smaller steps as we get close to the minimum and we exceed the threshold by a small margin compared to other methods which result in smaller norm of final gradient compared to threshold.

See the supplementary figures for question 1 for traces of the optimization process and plots for norm of the gradient across iterations.

2

For minimization problem arising from the following systems of equations:

$$Ax = 1,$$

where $A \in R^{n \times n}$ is a hilbert matrix, we obtain solutions for different dimenions of the hilbert matrix by using conjugate gradient method for minimization. The table below contains the number of iterations and corresponding condition numbers of the hilbert matrices. We can see that increasing the dimension results in more and more ill-conditioned matrices (i.e. larger values), which effectively increases the number of iterations to reach the desired threshold for the norm of the residual in conjugate gradient method.

See the supplementary figures for question 2 for the plot of the norm of the residuals across iterations.

n	iterations	condition number	final residual norm
5	6	4.7661e+05	5.5413e-09
8	19	1.5258e+10	5.8953e-09
12	38	1.6296e+16	6.4885e-07
20	63	2.9461e+18	6.5207e-07

3

To find the minimizer of Rosenbrock function

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

we use the trust region method, in which the sub-problem of minimization of the quadratic model is addressed by Steihaugh's linear Conjugate Gradient method, which does not require the B_k in the quadratic model to be a positive definite matrix.

Given the initial point $x_0 = (0, -1)^T$, we are able to react to the minimizer $x^* = (1, 1)^T$ with 23 iterations and final stopping tolerance is met. Comparing this iteration number to the ones obtained with steepest descent (160) and newton (4), it improves upon the steepest descent but lacks behind the newton's and quasi newton methods tried in the laboratory hour 1. This is due to the fact that trust region method used with conjugate gradient methods brings super-linear convergence. Newton and Quasi-Newton line search methods can converge quadratically but does not guarantee the convergence. However, in our problem, they converge and they do so faster than this trust region method.

See the supplementary figures for question 3 for the contour plots of quadratic models in every 5 iteration, and iteration trace of the optimization process.