Shashank Singh[1]

3 Newton and quasi-Newton methods [25 points] (Sashank)
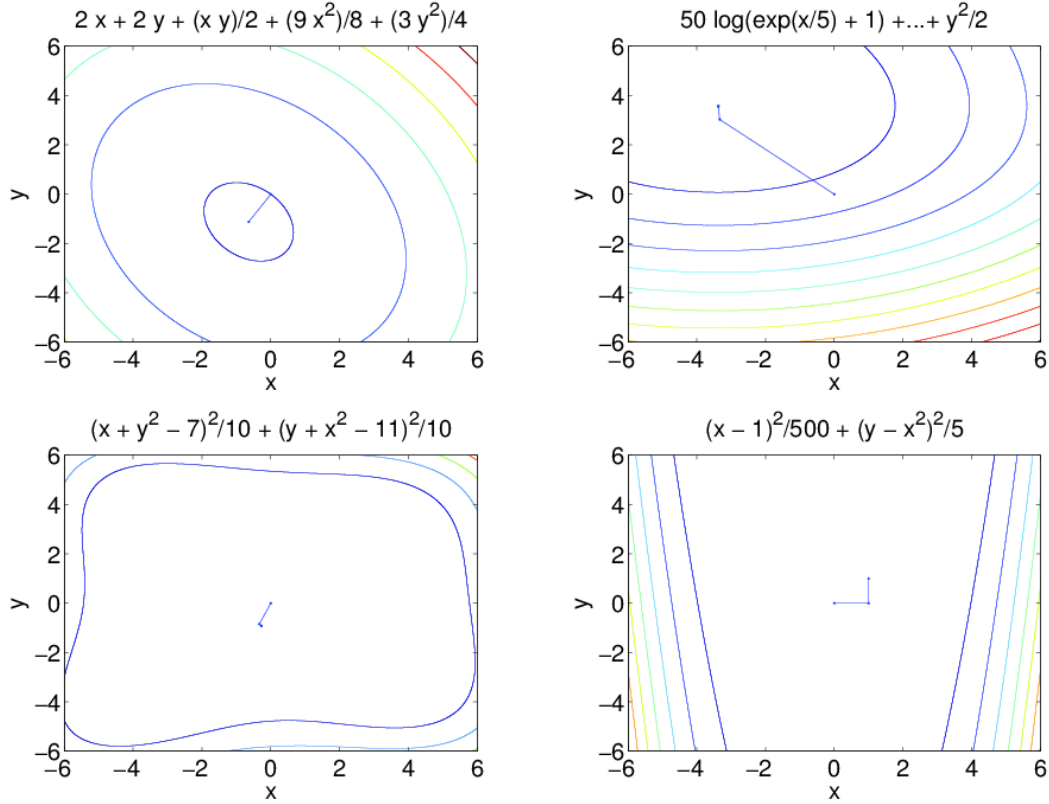
(a) See Figure 1.

Figure 1: Trajectories of Newton's Method (starting at $(0,0)$) and contour plots of each function.

(b) See Figure 2.

(c) The final points, and hence the function values, are the same for both algorithms (see Tables 1,2, and 3), except on f_H starting at $(2,2)$. This makes sense, because f_H is the most obviously non-convex of the 4 functions. Examining the trajectory of Newton's method in this case shows that it very quickly shoots off far from the closest minimum before converging to another minimum. Backtracking prevents this from happening with BFGS.

(d) On f_Q, f_{LL}, and f_H, the number of iterations for both algorithms are comparable (differing by at most 2; see Table 4). It's possible that the starting positions are simply too close to the optima to notice a difference. However, for f_R, BFGS consistently requires many more iterations than Newton's method. Iterations of Newton's method are generally $O(n^p)$, for some
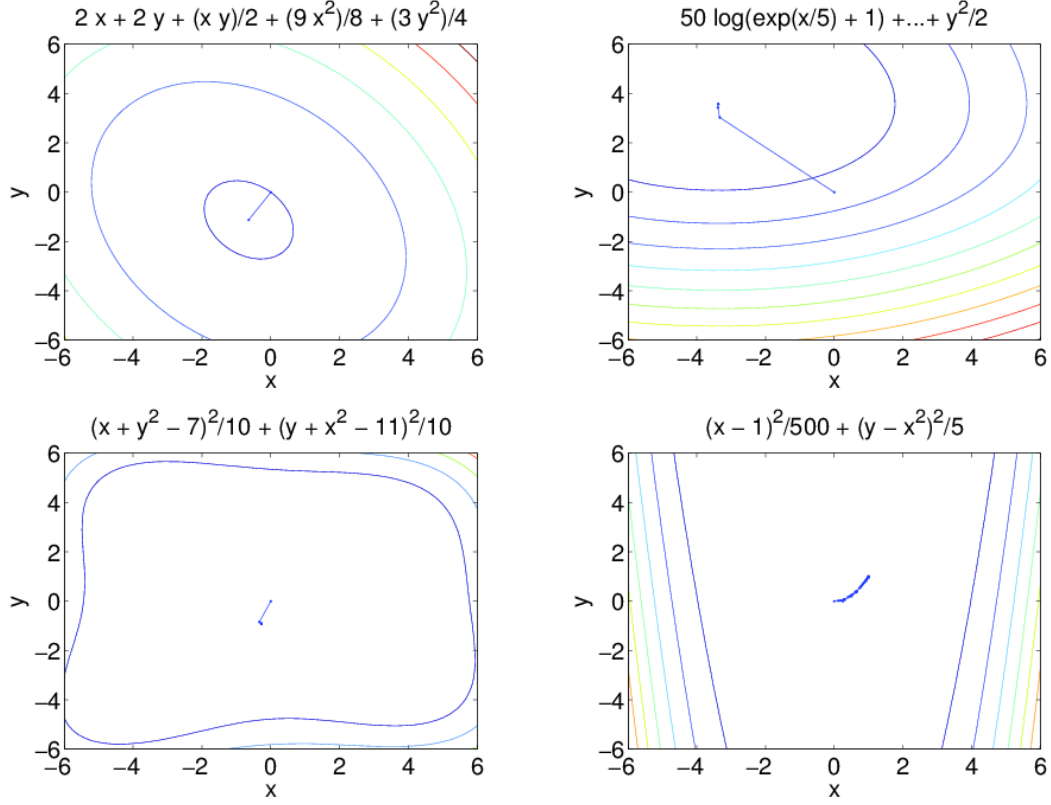
[1]sss1@andrew.cmu.edu

Figure 2: Trajectories of the BFGS algorithm (starting at $(0,0)$) and contour plots of each function.

$p \in [2.3, 3]$) (due to the inversion of the $n \times n$ Hessian). The most expensive operations in an iteration of BFGS are several multiplications of n-dimensional vectors by $n \times n$ matrices, requiring $O(n^2)$ time.

Function	Newton final point	BFGS final point	Newton function value	BFGS function value
f_Q	$(-0.6400, -1.1200)$	$(-0.6400, -1.1200)$	-1.7600	-1.7600
f_{LL}	$(-3.3742, 3.5787)$	$(-3.3742, 3.5787)$	40.4012	40.4012
f_H	$(-0.2708, -0.9230)$	$(-0.2708, -0.9230)$	18.1617	18.1617
f_R	$(1.0000, 1.0000)$	$(1.0000, 1.0000)$	0	0.0000

Table 1: Final values and functions values for each algorithm and function, starting from $(0,0)$.

Function	Newton final point	BFGS final point	Newton function value	BFGS function value
f_Q	$(-0.6400, -1.1200)$	$(-0.6400, -1.1200)$	-1.7600	-1.7600
f_{LL}	$(-3.3742, 3.5787)$	$(-3.3742, 3.5787)$	40.4012	40.4012
f_H	$(3.5844, -1.8481)$	$(3.0000, 2.0000)$	0.0000	0.0000
f_R	$(1.0000, 1.0000)$	$(1.0000, 1.0000)$	0	0.0000

Table 2: Final values and functions values for each algorithm and function, starting from $(2, 2)$.

Function	Newton final point	BFGS final point	Newton function value	BFGS function value
f_Q	$(-0.6400, -1.1200)$	$(-0.6400, -1.1200)$	-1.7600	-1.7600
f_{LL}	$(-3.3742, 3.5787)$	$(-3.3742, 3.5787)$	40.4012	40.4012
f_H	$(3.0000, 2.0000)$	$(3.0000, 2.0000)$	0.0000	0.0000
f_R	$(1.0000, 1.0000)$	$(1.0000, 1.0000)$	0	0.0000

Table 3: Final values and functions values for each algorithm and function, starting from $(3, 3)$.

Function	Newton iterations starting at $(0,0)$	BFGS iterations starting at $(0,0)$	Newton iterations starting at $(2,2)$	BFGS iteration starting at $(2,2)$
f_Q	2	2	2	2
f_{LL}	5	7	5	6
f_H	5	7	9	9
f_R	3	37	6	44
	Newton iterations starting at $(3,3)$	BFGS iterations starting at $(3,3)$		
f_Q	2	2		
f_{LL}	4	6		
f_H	6	8		
f_R	6	52		

Table 4: Number of iterations for each algorithm and function, at various starting points.