

3 Programming gradient descent [25 points] (Yifei)

- (a) (5 pts) Based on the plots, f_Q , f_{LL} , and f_R appear convex, whereas f_H appears to have minima that are local but not global.

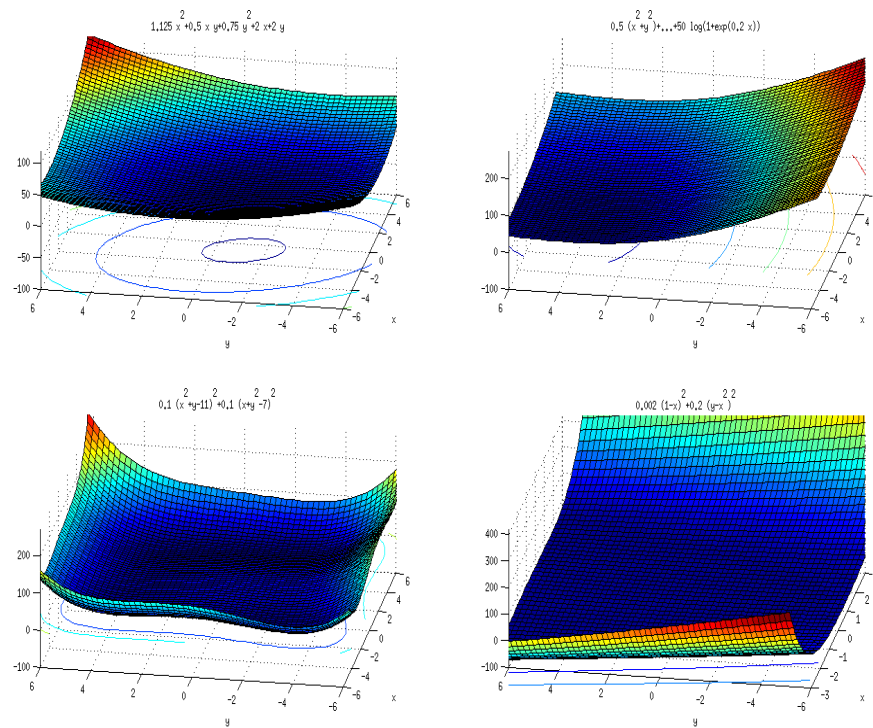


Figure 1: Surface and contour plots of the four objective functions.

- (b) (8 pts) In each figure below, the row indicates the step size (0.3 or 0.8) and the column indicates the initialization $((2, 3)^T$ or random).
- (c) (6 pts)
- (d) (4 pts)
- (e) (2 pts)

¹sssl@andrew.cmu.edu

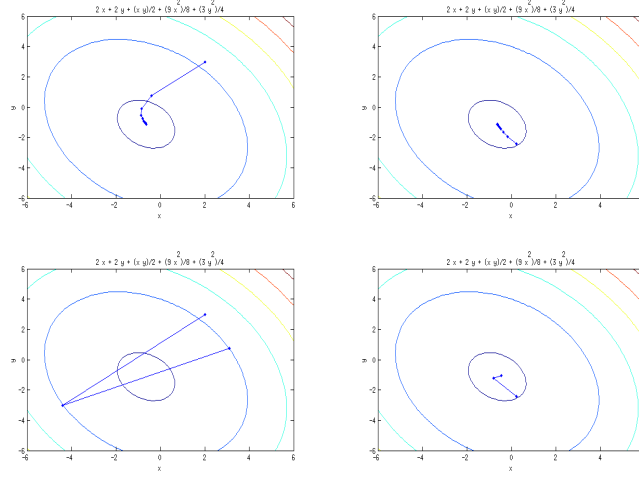


Figure 2: Gradient descent path and contour plots of f_Q at each step size and initialization.

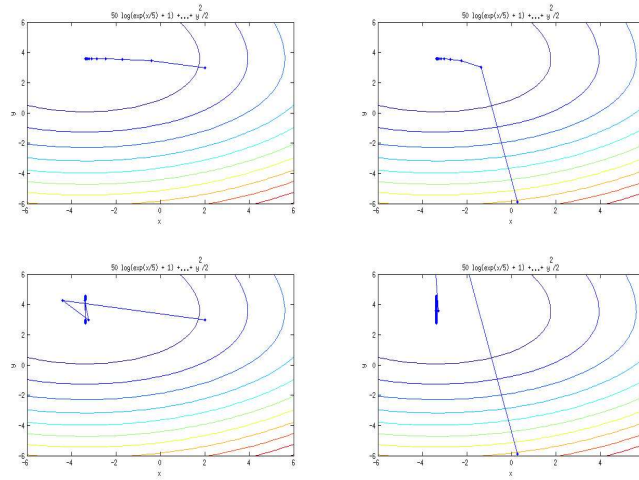


Figure 3: Gradient descent path and contour plots of f_{LL} at each step size and initialization.

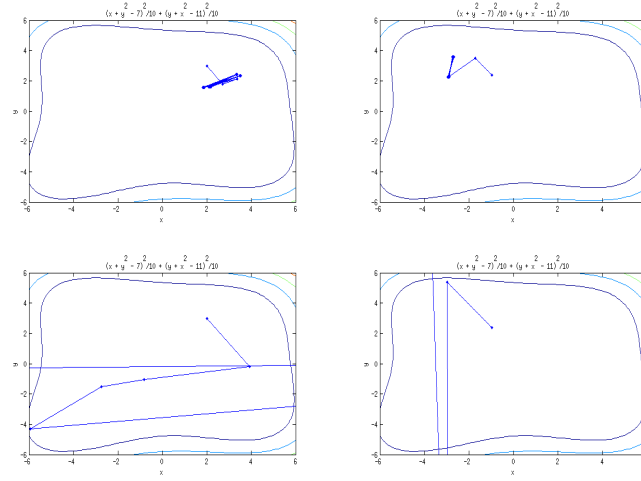


Figure 4: Gradient descent path and contour plots of f_H at each step size and initialization.

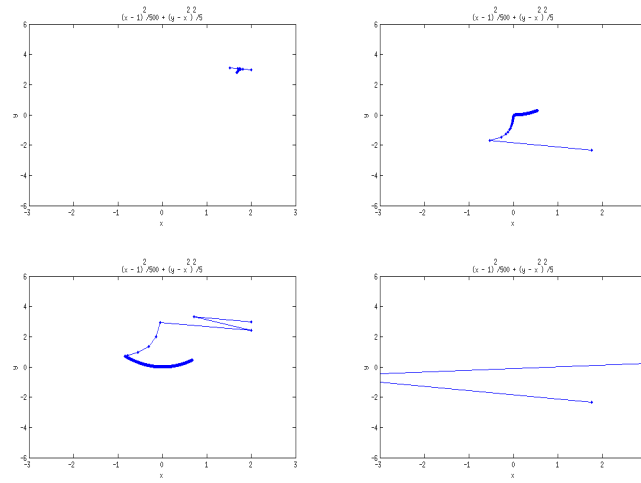


Figure 5: Gradient descent path and contour plots of f_R at each step size and initialization.