Problem 1.1.4

use optimality conditions to show that

for all x > 0 we have 1/x + x >= 2

$Let\ f(x) = \frac{1}{x} + x\ and\ domain\ of\ f = \{x \mid x \in R, x > 0\}$

The necessary optimality conditions are below.

$\nabla f(x) = 0\ and\ \nabla^2 f(x) \geq 0$

+ First, $\nabla f(x) = 0$

$-1x^{-2} + 1 = 0$

$-1 + x^2 = 0$

$x = \pm 1$

$x = 1\ because\ x > 0$

+ Second, $\nabla^2 f(x) \geq 0$

$2x^{-3} \geq 0$

And plug in x = 1.

$2 \geq 0$

+ Afterall, global min of f(x) is at x = 1 and f(1) = 2.

$Therefore,\ \forall x > 0,\ f(x) \geq 2$

Problem 1.2.1

1.2.1 minimizing $f(x, y) = 3x^2 + y^4$

Applying one iteration of steepest descent method with Armijo rule $s = 1,\ \sigma = 0.1,\ \beta = 0.5$

· Armijo Rule

$$f(x^k) - f(x^k + \beta^m \cdot s \cdot d^k) \geq -\sigma \beta^m s \nabla f(x^k)' d^k$$

$$f(x^k) + \sigma \beta^m s \nabla f(x^k)' d^k \geq f(x^k + \beta^m s d^k) \quad \cdots\cdots (*)$$

⟨(a) solution⟩

· Applying one iteration $k=0$ and we're going to find m until the above $*$ inequality is satisfied.

$$f(x^0) + 0.1(0.5)^m \nabla f(x^0)' d^0 \geq f(x^0 + (0.5)^m d^0) \cdots (**)$$

· $d^k = -\nabla f(x^k) = -\begin{bmatrix} 6x^k \\ 4(y^k)^3 \end{bmatrix}$

$$d^0 = -\begin{bmatrix} 6x^0 \\ 4(y^0)^3 \end{bmatrix} = -\begin{bmatrix} 6 \cdot 1 \\ 4(-2)^3 \end{bmatrix} = \begin{bmatrix} -6 \\ 32 \end{bmatrix}$$

$$\nabla f(x^0)' = [6 \quad -32]$$

$$f(x^0, y^0) = 3 \cdot 1^2 + (-2)^4 = 3 + 16 = 19$$

· Now for $m=0$, the above $(**)$ is

$$LHS = 19 + 0.1(0.5)^0 [6 \quad -32] \begin{bmatrix} -6 \\ 32 \end{bmatrix}$$

$$= 19 + 0.1 \times (-36 - 1024) = 19 - 106 = -87$$

$$RHS = f(x^0 + (0.5)^m d^0) = f\left(\begin{bmatrix} x^0 \\ y^0 \end{bmatrix} - (0.5)^m \begin{bmatrix} 6x^0 \\ 4(y^0)^3 \end{bmatrix} \right)$$

$$= f\left(\begin{bmatrix} x^0 - (0.5)^m \cdot 6x^0 \\ y^0 - (0.5)^m \cdot 4 \cdot y^{0^3} \end{bmatrix} \right) = f\begin{pmatrix} 1 - (0.5)^m \cdot 6 \\ -2 + (0.5)^m \cdot 32 \end{pmatrix}$$

$$= 3(1 - 6(0.5)^m)^2 + (-2 + 32(0.5)^m)^4 = 3(-5)^2 + (-2+32)^4 = 810075$$

so, $-87 \not\leq 810075$

Armijo condition isn't met.

- For $m=1$, the above ** is

$LHS = 19 + 0.1(0.5)^1 [6 \ -32] \begin{bmatrix} -6 \\ 32 \end{bmatrix}$

$= 19 + 0.05(-1060) = 19 - 53 = -34$

$RHS = 3(1-6(0.5))^2 + (-2+32(0.5))^4$

$= 3(1-3)^2 + (-2+16)^4$

$= 12 + 38416 = 38428$

so, $-34 \not\leq 38428$

Armijo condition isn't met.

- For $m=2$, the above ** is

$LHS = 19 + 0.1(0.5)^2 [6 \ -32] \begin{bmatrix} -6 \\ 32 \end{bmatrix}$

$= -7.5$

$RHS = 3(1-6(0.5)^2)^2 + (-2+32(0.5)^2)^4$

$= 1296.75$

so, $-7.5 \not\leq 1296.75$. Armijo condition is not met.

- For $m=3$, the above ** is

$LHS = 19 + 0.1(0.5)^3 [6 \ -32] \begin{bmatrix} -6 \\ 32 \end{bmatrix} = 5.75$

$RHS = 3(1-6(0.5)^3)^2 + (-2+32(0.5)^3)^4$

$= 16.1875$

so, $5.75 \not\leq 16.1875$. Armijo condition isn't met

· For $m=4$

$LHS = 19 + 0.1(0.5)^4 [6\ -32] \begin{bmatrix} -6 \\ 32 \end{bmatrix} = 12.375$

$RHS = 3(1-6(0.5)^4)^2 + (-2+32(0.5)^4)^4$
$= 3\cdot(0.625)^2 + 0^4 = 1.171875$

so, $12.375 \geq 1.171875$

Armijo condition is met.

Therefore, $m=4$ and we have $(0.625, 0)$ for the next iteration.

((b solution))

$s=1,\ \sigma=0.1,\ \beta=0.1$

Everything is the same, except the reduction factor

For $m=0$, Armijo condition isn't met because it's the same result like previous one

For $m=1$,

$LHS = 19 + 0.1(0.1)^1 [6\ -32] \begin{bmatrix} -6 \\ 32 \end{bmatrix} = 8.4$

$RHS = 3(1-6(0.1)^1)^2 + (-2+32(0.1)^1)^4$
$= 3(0.4)^2 + (1.2)^4 = 2.5536$

so, $8.4 \geq 2.5536$

Armijo condition is met.

Therefore, $m=1$ and we have $(0.4, 1.2)$

I think the cost of the new iteration is low because of lower β, reduction factor.
We can reach the proper m fast.

(c solution)

- newton's method when stepsize = 1 & starting point (1, -2)

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)}$$

- $f(x,y) = 3x^2 + y^4$

- $\frac{d}{dx} f(x,y) = 6x$ $\qquad$ $\frac{d}{dy} f(x,y) = 4y^3$

- 1 iteration of x

$$x' = x^0 - \frac{f(x^0)}{f'(x^0)}$$

$f(x^0) = f(1) = 3 \cdot 1^2 = 3$

$f'(x^0) = f'(1) = 6 \cdot 1 = 6$

$x' = 1 - \frac{3}{6} = 0.5$

- 1st iteration of y

$$y' = y^0 - \frac{f(y^0)}{f'(y^0)}$$

$f(y^0) = f(-2) = (-2)^4 = 16$

$f'(y^0) = f'(-2) = 4 \cdot (-2)^3 = -32$

$y' = -2 - \frac{16}{-32} = -2 + 0.5 = -1.5$

(a) Newton's method only took a step getting a result compare to part (a) which took five steps ($m=4$.)

(b) And amount of work of Newton's method is quite simple and fast.

3.1.b.1

L-BFGS-B is an optimization algorithm using quasi-Newton methods based on a limited amount of computer memory. The goal problem of the algorithm is to minimize the unconstrained values of a real vector x, where f(x) is a differentiable scalar function. It is a widely used algorithm for parameter estimation in machine learning.

3.1.b.2

filename	_total_solve_time	Optimal value
dqrtic.mod	0.087309	1.99817e-14
eigenbls.mod	0.076656	3.52907e-08
freuroth.mod	0.161427	608159

3.2.b

n value	_total_solve_time	Optimal value
10	0.001721	80.5912
20	0.001787	173.122
50	0.001507	396.986
100	0.00195	777.058
1000	0.010822	8644.15
10000	0.095459	87993.6

+ Compare their performance and discuss your observation.

- In terms of time, I think the time grows as n value grows and it's not exactly linear, but I can say it's not like exponential. The time complexity is in between linear and exponential and I looked up internet and I found out that Rastrigin follows nln(n).

- In terms of optimal value, the values are almost 80 times with n values. The results are consistent.