Chapter 5 Exercises

Question 5.3

(a) The derivative of $f(x) = x^2 - y$ is $f'(x) = 2x$, therefore the Newton iteration for solving $f(x) = 0$ is $x_{k+1} = x_k - (x_k^2 - y)/2x_k$.

(b) The general rule for getting x bits of accuracy given a 4 bit accurate initial guess is to solve for the number of iterations k in $4 * 2^k = x$. To get 24 bits:

$$k = log_2(24/4) = 3$$

so we need 3 iterations.

For 53 bits we need

$$k = log_2(53/4) = 3.72$$

which rounds up to 4 iterations.

Question 5.4

We take the given function

$$f(x) = x - 1 - y$$

and take its derivative:

$$f'(x) = -x - 2.$$

From this, the Newton iteration for solving $f(x) = 0$ is given by

$$x_{k+1} = x_k - (x_k^{-1} - y)/(-x_k^{-2}) = x_k + x_k^2(x_k^{-1} - y) = 2x_k - x_k^2 y.$$

Since we reduced the terms, there are no divisions in the final formula.

Question 5.6

(a) Since $g_1'(x) = 1 - 2x$ and $|g_1'(\sqrt{3})| = |1 - 2\sqrt{3}| \approx 2.46 > 1$, the iterative scheme is not convergent.

(b) Since $g_2'(x) = 1 - 2x/y$ and $|g_1'(\sqrt{3})| = |1 - 2\sqrt{3}/3| \approx 0.155 < 1$ so the iterative scheme is locally convergent.

(c) $f'(x) = 2x$, so the fixed point iteration function given by Netwon's method is $g(x) = x - (x^2 - y)/2x$.

Question 5.9

$x_1^2 - x_2^2 = 0$, and $2x_1x_2 = 1$ that is

$$f(x) = \begin{bmatrix} x_1^2 - x_2^2 \\ 2x_1x_2 - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

with starting value $x_0 = \begin{pmatrix} 0 & 1 \end{pmatrix}^T$.

Our Jacobian is given by

$$J_f(x) = \begin{pmatrix} 2x_1 & -2x_2 \\ 2x_2 & 2x_1 \end{pmatrix}.$$

Since we were given a starting vector of $x_0 = [0, 1]^T$ we have

$$f(x_0) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

and

$$J_f(x_0) = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$$

solving the system

$$\begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix} s_0 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

gives us

$$s_0 = \begin{bmatrix} -1/2 & 1/2 \end{bmatrix}^T$$

$$x_1 = x_0 + s_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}.$$

The result of the first iteration of the Newton method is

$$x_1 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}.$$

Question 5.10

If $x_k = x*$ then $f(x_k) = 0$ and

$$x_{k+1} = x_k - f(x_k)\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} = x * -0 = x * .$$

If $x_{k-1} = x*$, then $f(x_{k-1}) = f(x*) = 0$ and

$$x_{k+1} = x_k - f(x_k)\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} = x_k - f(x_k)\frac{x_k - x*}{f(x_k)} = x_k - (x_k - x*) = x*.$$

in both cases, the next value of x is $x*$.

Chapter 5 Computer Exercises

Check the Jupyter notebook! :)

Chapter 6 Exercises

Question 6.1

(a)

$$f(x) = \frac{1}{2}(x_1^2 - x_2)^2 + \frac{1}{2}(1 - x_1)^2.$$

Both terms are positive, so the smallest value we can hope for is 0. The second term is zero when $x_1 = 1$, if we hold this then the first term is zero when $x_2 = 1$. The Function therefore achieves a zero at $\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$.

(b)

$$x_1 = x_0 - H_f^{-1}(x_k)\Delta f(x_k).$$

$$\Delta f(x) = \begin{bmatrix} 2x_1^3 - 2x_2x_1 + x_1 - 1 \\ -x_1^2 + x_2 \end{bmatrix} = \left[\begin{smallmatrix} 9 \\ -2 \end{smallmatrix}\right]$$

$$H_f(x) = \begin{bmatrix} 6x_1^2 - 2x_2 + 1 & -2x_1 \\ -2x_1 & 1 \end{bmatrix} = \left[\begin{smallmatrix} 21 & -4 \\ -4 & 1 \end{smallmatrix}\right]$$

The newton equation to be solved is

$$\left[\begin{smallmatrix} 21 & -4 \\ -4 & 1 \end{smallmatrix}\right] s_0 = \left[\begin{smallmatrix} 9 \\ -2 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0.2 \\ -1.2 \end{smallmatrix}\right]$$

and the next iteration value is

$$x_1 = x_0 + s_0 = \left[\begin{smallmatrix} 2 \\ 2 \end{smallmatrix}\right] + \left[\begin{smallmatrix} 0.2 \\ -1.2 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 2.2 \\ 0.8 \end{smallmatrix}\right]$$

(c) Our result gets much closer to the correct value for the x_2 term, however *(d)* our result gets slightly farther away for the x_1 term.

Question 6.2

The gradient and Hessian are given by:

$$f(x) = \frac{1}{2} * \left[x_1 * (A_{11}x_1 + A_{12}x_2 + \ldots) + x_2 * (A_{21}x_1 + A22x_2 + \ldots) + \ldots + x_i(Ai1x_1 + \ldots + Aiix_i) \right]$$

$$\Delta f(x) = \frac{1}{2} * \begin{bmatrix} A_{11} * (A_{11}x_1 + A_{12}x_2 + \ldots) + A_{21} * x_2 + \ldots + A_{i1} * x_i \\ A_{22} * (A_{21}x_1 + A_{22}x_2 + \ldots) + A_{12} * x_1 + \ldots + A_{i2} * x_i \\ \ldots \\ A_{ii} * (A_{i1}x_1 + A_{i2}x_2 + \ldots) + A_{1i} * x_1 + \ldots + A_{i-1,i} * x_{i-1} \end{bmatrix}$$

$$= \frac{1}{2} * \begin{bmatrix} A_{11}^2 x_1 + (A_{11}+1)*(A_{12}x_2 + A_{1i}x_i) \\ A_{22}^2 x_2 + (A_{22}+1)*(A_{21}x_1 + ... + A_{2i}x_i) \\ ... \\ A_{ii}^2 x_i + (A_{ii}+1)*(A_{i1}x_1 + A_{i2}x_2 + ...) \end{bmatrix}$$

Damn, I don't know what to do... My guess is that the b vector subtracts out the ugly terms. When we do the hessian we're left with like the diagonal or something.

I don't know what argument to make about what the minimal value should be though, like should it be when the f(x) = 0? I think I'm just going to skip this question.

Question 6.3

From the discussion in section 6.5, we can see clearly that the Hessian of the Lagrangian function is

$$H_L(x, \lambda) = \begin{bmatrix} B(x,\lambda) & J_g^T(x) \\ J_g(x) & O \end{bmatrix}$$

Let's look at the 2x2 case with the Hessian as

$$\begin{bmatrix} B & J^T \\ J & 0 \end{bmatrix}$$

for a vector (x, y) we can set up the following:

$$\begin{pmatrix} x^T & y^T \end{pmatrix} \begin{pmatrix} B & J^T \\ J & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x^T B x + 2y^T J x$$

If $x = 0$ and we pick any value of y, the result equals zero, therefore we pass the test for a positive definite matrix and the matrix can not be positive definite.

Question 6.4

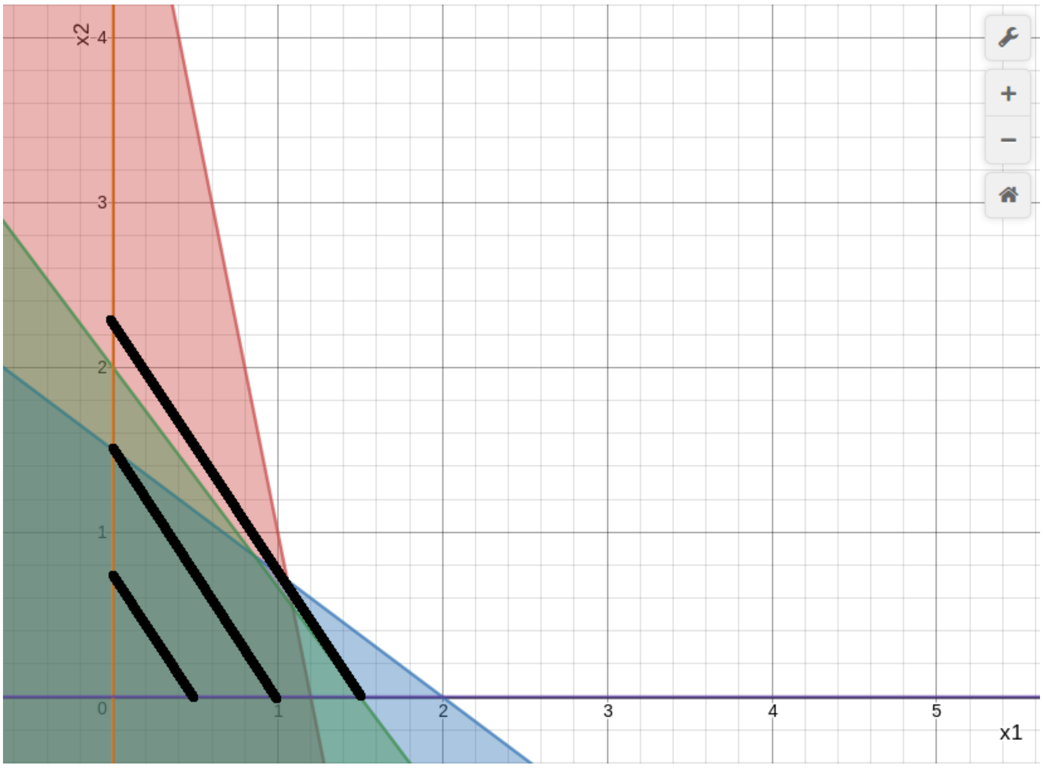
There is one at $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, and the on-axis ones are $\begin{bmatrix} 1.2 \\ 0 \end{bmatrix} and \begin{bmatrix} 0 \\ 1.5 \end{bmatrix}$, the other three points are given by the intersections of the three lines, which occur at $\begin{bmatrix} \frac{18}{17} \\ \frac{12}{17} \end{bmatrix}$, $\begin{bmatrix} \frac{6}{7} \\ \frac{6}{7} \end{bmatrix}$ and $\begin{bmatrix} \frac{12}{11} \\ \frac{6}{11} \end{bmatrix}$.

Of which it is clear that

$$\begin{bmatrix} \frac{18}{17} \\ \frac{12}{17} \end{bmatrix}$$

is dominated by the other two points. Therefore we have a region with 5 vertices:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1.2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1.5 \end{bmatrix}, \begin{bmatrix} \frac{6}{7} \\ \frac{6}{7} \end{bmatrix}, \begin{bmatrix} \frac{12}{11} \\ \frac{6}{11} \end{bmatrix}.$$



Evaluated at each of these points the result is (in the same order):

$$0.0, -3.6, -3.0, -4.28, -4.36$$

Therefore the minimum value occurs at

$$\begin{bmatrix} \frac{12}{11} \\ \frac{6}{11} \end{bmatrix}.$$

The graph is shown in the following figure

Chapter 6 Computer Exercises

Check the Jupyter notebook! :)