Introduction to Optimization

Fall 2015, Homework 1

1. The gradient and Hessian matrix are, respectively,

$$\nabla f(x) = \begin{bmatrix} -1/2 x_1^{-1/2} x_2^{1/2} \\ -1/2 x_1^{1/2} x_2^{-1/2} \end{bmatrix}, \quad H_f(x) = \begin{bmatrix} 1/4 x_1^{-3/2} x_2^{1/2} & -1/4 x_1^{-1/2} x_2^{-1/2} \\ -1/4 x_1^{-1/2} x_2^{-1/2} & 1/4 x_1^{1/2} x_2^{-3/2} \end{bmatrix}.$$

Now,

$$\nabla f(x)^\top y = -\frac{1}{2} x_1^{-1/2} x_2^{1/2} y_1 - \frac{1}{2} x_1^{1/2} x_2^{-1/2} y_2, \tag{0.1}$$

$$y^\top H_f(x) y = \frac{1}{4}\left(x_1^{-3/2} x_2^{1/2} y_1^2 - 2 x_1^{-1/2} x_2^{-1/2} y_1 y_2 + x_1^{1/2} x_2^{-3/2} y_2^2 \right) \tag{0.2}$$

$$= \frac{1}{4} x_1^{1/2} x_2^{1/2} \left(\frac{y_1}{x_1} - \frac{y_2}{x_2} \right)^2$$

From (0.2), one sees that the Hessian matrix is positive semi-definite, and therefore f is convex. However, f is not strictly convex, since (0.2) equals 0 for $\frac{y_1}{x_1} = \frac{y_2}{x_2}$.

2. For the Newton method we need the first and second derivatives:

$$f'(x) = -\frac{1}{2}\sin(2x) + x\,\mathrm{e}^{-\frac{x^2}{2}},$$

$$f''(x) = -\cos(2x) + \mathrm{e}^{-\frac{x^2}{2}} - x^2\,\mathrm{e}^{-\frac{x^2}{2}}.$$

Now

$$\begin{aligned} x_0 &= 1 \text{ (this was given)}, \\ x_1 &\approx 0.6350, \\ x_2 &\approx 0.4183, \\ &\vdots \\ x_{44} &\approx 2.488 * 10^{-8}. \end{aligned}$$

For x_{44}, the gradient equals 0.
For the fixed point iteration, we obtain

$$\begin{aligned} x_0 &= 1, \\ x_1 &\approx -11.5327, \\ x_2 &\approx 0.0793, \\ x_3 &\approx -0.0007, \\ x_4 &\approx 6.2415 * 10^{-10}, \\ x_5 &\approx 0.0000. \end{aligned}$$

For x_5, $g(x_5) = 0$.
The fixed point iteration is now much faster. For the fixed point we need five iterations but for the Newton we need 44 iterations. The reason that the Newton method is slow, is that the minimum point is the same where the $f''(x) = 0$ (that is the denominator in the method). Thus more we approach the minimum, more singular is the next iteration. The fixed point iterations converges if $|g'(x)| < 1$ whenever x is near the minimum $x_{opt} = 0$. This really holds in the interval $(-0.46, 0.46)$.