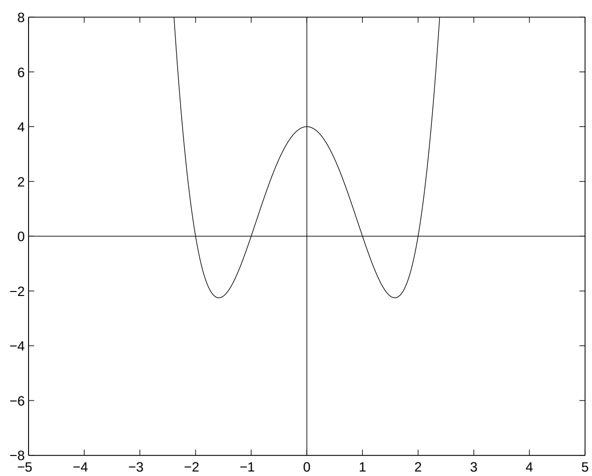
Solutions for Homework 9 Foundations of Computational Math 1 Fall 2011

Problem 9.1

Let $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ be given by

$$f(x) = x^4 - 5x^2 + 4$$

and consider applying Newton's method for optimization. Here Newton's method refers to the basic form where the step size is 1 and nothing is done to alter the Hessian to guarantee positive definiteness. Note that $f(x)$ is a scalar function of a scalar argument and has the form



(i) What are the values of x that are local minimizers or local maximizers of $f(x)$. Justify your answers.

(ii) Find the value $\beta > 0$ such that $f(x)$ has negative curvature for $-\beta < x < \beta$, and positive curvature outside the interval, i.e., for $x < -\beta$ or $x > \beta$.

(iii) What happens to the Newton step at $x = \beta$?

(iv) Determine $\mu(x) : \mathbb{R} \rightarrow \mathbb{R}$ such that the step of Newton's method applied to $f(x)$ can be written as $x_{k+1} = \mu(x_k)x_k$.

(v) Find the value of $\alpha \in \mathbb{R}$ such that $\beta > \alpha > 0$ and Newton's method cycles and does not converge when $x_0 = \alpha$ or $x_0 = -\alpha$. That is, $-\alpha = \mu(\alpha)\alpha$ and $\alpha = -\mu(-\alpha)\alpha$.

(vi) Show that if $-\alpha < x < \alpha$ then

$$|\mu(x)| < 1$$

(vii) Show that if $-\alpha < x_0 < \alpha$ is the initial point for Newton's method then there is a constant $0 < \gamma < 1$ (possibly dependent on x_0 but independent of k) such that

$$|x_{k+1}| < \gamma |x_k|$$

and therefore $x_k \to 0$.

(viii) It can be shown that if $x_0 > \sigma > 0$ where σ is the rightmost local minimizer of $f(x)$ then $x_k \to \sigma$ for Newton's method. Using this fact and those above, show that it is possible to choose $-\beta < x_0 < -\alpha$ so that $x_k \to \sigma$ for Newton's method.

(ix) Implement Newton's method (in MATLAB preferably) and demonstrate the convergence behavior determined above.

Solution:

The list of questions posed are one way of analyzing the behavior of Newton's method when optimizing this, fairly simple, cost function. It is certainly not the only approach as can be seen from the variety of analyses of convergence behavior we have seen in the notes, textbook, and homework problems for nonlinear system solving and optimization.

We have

$$f = x^4 - 5x^2 + 4 \quad f' = 4x^3 - 10x = 2x(2x^2 - 5)$$
$$f'' = 12x^2 - 10 = 2(6x^2 - 5) \quad f''' = 24x$$
$$f'''' = 24$$

Since f is a scalar real function the gradient and Hessian are both scalar functions, i.e., f' and f'' respectively.

The critical points are the roots of $f' = 2x(2x^2 - 5)$ which are $\sigma_- = -\sqrt{5/2}$, $\sigma_+ = \sqrt{5/2}$, and $\sigma_0 = 0$. Evaluating $f'' = 2(6x^2 - 5)$ at these points yields $f''(\sigma_-) = 20$, $f''(\sigma_+) = 20$, and $f''(\sigma_0) = -10$. Therefore, σ_- and σ_+ are local minimizers and σ_0 is a local maximizer.

Negative curvature corresponds to the interval $-\beta < x < \beta$ where the Hessian is negative definite, i.e., $f''(x) < 0$. This implies that both possible directions $+1$ or -1 are directions of negative curvature in this interval.

$$f'' = 2(6x^2 - 5) < 0 \Leftrightarrow -\sqrt{5/6} < x < \sqrt{5/6}$$
$$\therefore \beta = \sqrt{5/6}$$

It is easily seen that since f'' is a quadratic polynomial it is positive outside the interval $-\beta < x < \beta$ as desired.

The Newton update to any x is $d(x) = -f'(x)/f''(x)$. At $x = \pm\beta$ we have $f'' = \mp\infty$ and therefore the Newton step is undefined.

The Newton step can be manipulated to yield $\mu(x)$ as follows:

$$x_+ = x + d(x) = x - \frac{f'(x)}{f''(x)} = \{\frac{4x^2}{6x^2 - 5}\}x = \mu(x)x$$

The condition $-\alpha = \mu(\alpha)\alpha \rightarrow \mu(\alpha) = -1$. So does the condition $\alpha = -\mu(-\alpha)\alpha$. It follows that

$$\mu(\alpha) = -1 \Leftrightarrow 4\alpha^2 = 5 - 6\alpha^2$$
$$\therefore \alpha = \sqrt{1/2}$$

If $-\alpha < x < \alpha$ then $6x^2 - 5 < 0$ and

$$|\mu(x)| = \frac{4x^2}{5 - 6x^2}$$
$$\frac{4x^2}{5 - 6x^2} < 1 \Leftrightarrow -\sqrt{1/2} < x < \sqrt{1/2}$$

as desired.

Clearly, if $-\alpha < x_0 < \alpha$ then x_k is a sequence with alternating sign and decreasing magnitude. To show that the limit is $\sigma_0 = 0$, i.e., the local maximizer, we need $|\mu(x_k)|$ to be nonincreasing as $x_k \rightarrow 0$ from either side. If this is true then we know that the contraction on the magnitude of x_k has been reduced by at least $|\mu(x_0)|^k$ and therefore converges to $\sigma_0 = 0$.

Consider $|\mu(x_k)|'$ on $-\sqrt{1/2} < x < \sqrt{1/2}$

$$|\mu(x)|' = \left[\frac{4x^2}{5 - 6x^2}\right]' = \frac{(5 - 6x^2)(8x) + (4x^2)(12x)}{(5 - 6x^2)^2}$$

The denominator is always positive on $-\sqrt{1/2} < x < \sqrt{1/2}$ so we need only consider the sign of the numerator.

$$0 < x < \sqrt{1/2} \rightarrow (5 - 6x^2)(8x) + (4x^2)(12x) > 0 \rightarrow |\mu(x)|' > 0$$
$$\therefore |\mu(x)| \rightarrow 0 \text{ from the right}$$
$$-\sqrt{1/2} < x < 0 \rightarrow (5 - 6x^2)(8x) + (4x^2)(12x) < 0 \rightarrow |\mu(x)|' < 0$$
$$\therefore |\mu(x)| \rightarrow 0 \text{ from the left}$$

We therefore have

$$|x_{k+1}| = |\mu(x_k)||x_k| < |\mu(x_0)||x_k| = \gamma|x_k|$$

as desired. It follow that if $-\sqrt{1/2} < x_0 < \sqrt{1/2}$ then $x_k \rightarrow 0 = \sigma_0$ the local maximizer.

Accepting the fact that if $x_0 > \sigma_+$ then $x_k \rightarrow \sigma_+$ we can find an initial condition $-\beta < x_0 < -\alpha$ so that $x_k \rightarrow \sigma_+$ by considering what happens near $-\beta$. Recall, $x_0 = -\beta$ has an undefined Newton step. More precisely, it has a step that positive and infinite. So if we consider $x_0 = -\beta + \epsilon$ with $\epsilon > 0$, by continuity, the step can be made positive and large enough to put $x_1 >> \sigma_+$. Convergence to σ_+ follows. This is easily verified via a MATLAB code. We have $-\beta \approx 0.9128$ and taking $x_0 = -0.85$ should yield convergence to $\sigma_+ = \sqrt{5/2}$.