Convergence of Levenberg-Marquardt (Part 1: Theory)

10 points

In this problem, we will encounter a function that demonstrates the value of the Levenberg-Marquardt method as an improvement over Gauss-Newton. That is, we will show, using the criteria for fixed point iterations that we saw in class, that Gauss-Newton will not converge, but Levenberg-Marquardt will, for this particular example. In part 2 of the problem (on the next page), we will verify this computationally.

Consider a function

$$F : \mathbb{R} \to \mathbb{R}^2 \qquad F(x) := \begin{bmatrix} x+1 \\ \lambda x^2 + x - 1 \end{bmatrix},$$

where λ is a real scalar parameter.

1. Show that the nonlinear least squares problem

$$\Phi(x) := \frac{1}{2}\|F(x)\|_2^2$$

with $\lambda < 1$ has a local minimum at $x = 0$. Also, show that for $\lambda < \frac{7}{16}$, the only real local minimum is at $x = 0$.

2. Prove that if $\lambda < -1$, $x = 0$ is a repelling fixed point of the Gauss-Newton method. In other words, show that $|g'(x=0)| > 1$ when $\lambda < -1$. $g(x)$ is the fixed point problem associated with the Gauss-Newton method.

3. Show that, on the other hand, the iterations of Levenberg-Marquardt method for $\lambda < \frac{7}{16}$ and $\mu = |\lambda|$ will converge. In other words, show that $|g'(x=0)| < 1$. $g(x)$ is the fixed point problem associated with the Levenberg-Marquardt method.

Be sure to show all your work and provide justifications for every step to receive full credit.

Please submit your response to this written problem as a PDF file below. You may do either of the following:

- write your response out by hand, scan it, and upload it as a PDF.

  We will not accept unprocessed pictures taken with your phone.

  If you decide to use your phone for scanning, make sure to use an app such as CamScanner (https://www.camscanner.com/) to get a readable PDF. Alternatively, there's a fast and convenient scanner in the Engineering IT office in 2302 Siebel that can just email you a PDF. (It's the Fax-machine-looking thing--not the scanner that's attached to one of the computers.)

- create the PDF using software.

If you're looking for an easy-ish way to type math, check out TeXmacs (http://texmacs.org/) or LyX (http://www.lyx.org/). Both are installed in the virtual machine. (Under "Applications / Accessories / GNU TeXmacs editor" and "Applications / Office / LyX document processor" respectively.)

Submit your response to each problems in this homework as a separate PDF. If you have multiple PDFs that you need to merge into one, try PDF Split and Merge (http://www.pdfsam.org/download/).

NOTE: Please make sure your solutions are legible and easy to follow. If they are not, we may deduct up to five points *per problem*.

Review uploaded file (blob:https://relate.cs.illinois.edu/02b4982e-8b15-416e-ba96-3551041195c4) · Embed viewer

Uploaded file*

选择文件 未选择任何文件

Your answer is correct.

1. We want to find minimum of $\Phi(x)$. The first thing we can note is that

$$\|F(x)\|_2 = \sqrt{(x+1)^2 + (\lambda x^2 + x - 1)^2} = \sqrt{\lambda^2 x^4 + 2\lambda x^2 + 2x - 2\lambda x}$$

and

$$\Phi(x) = \frac{1}{2}\lambda^2 x^4 + \lambda x^3 + (1-\lambda)x^2 + 1.$$

$$\Phi'(x) = 2\lambda x^3 + 3\lambda x^2 + (2-2\lambda)x = 0$$

From here, we see that $x = 0$ is a critical point. Now, we need to classify this critical point to verify that it is a minimum.

$$\Phi''(x=0) = 2 - 2\lambda$$

In order for this to be a minimum, $2 - 2\lambda > 0$. This implies that $\lambda < 1$. Thus, if $\lambda < 1$, the nonlinear compensation problem has a local minimum at $x = 0$.

In order for $x = 0$ to be the only real local minimum, we need to show that the remainder of the roots of $\Phi(x)$ are complex. The remainder of the roots are associated with $2\lambda^2 x^2 + 3\lambda x + (2 - 2\lambda)$. Since this is a second degree polynomial, we can find its roots using the quadratic formula. Therefore, if we can show that the discriminant is less than 0, then the roots will be complex. Discriminant: $9\lambda^2 - 16\lambda^2 + 16\lambda^3$.

$$\begin{aligned} 9\lambda^2 - 16\lambda^2 + 16\lambda^3 &< 0 \\ \Rightarrow 9 + 16\lambda &< 16 \\ \Rightarrow \lambda &< \frac{7}{16} \end{aligned}$$

Thus, when $\lambda < \frac{7}{16}$, the roots are 0 or complex and the only real local minimum is at $x = 0$.

2. The fixed point problem is

$$g(x) = x - \left(\boldsymbol{J}_{\boldsymbol{r}}^T(x)\,\boldsymbol{J}_{\boldsymbol{r}}(x)\right)^{-1}\boldsymbol{J}_{\boldsymbol{r}}^T(x)\,\boldsymbol{r}(x),$$

where $r(x) = F(x)$. In order to show that $x = 0$ is a repelling fixed point, we need to show that $|g(x^*)| > 1$ when $\lambda < -1$.

Since we have a specific function F, we can more explicitly compute the form of $g(x)$.

$$\begin{aligned} g(x) &= x - \left[\frac{(x+1) + (2\lambda x+1)(\lambda x^2 + x - 1)}{1 + (2\lambda x + 1)^2}\right] \\ &= x - \left[\frac{2\lambda^2 x^3 + 3\lambda x^2 + (2-2\lambda)x}{4\lambda^2 x^2 + 4\lambda x + 2}\right] \end{aligned}$$

We can now more readily compute its derivative using the quotient rule. We can also simplify this a bit further knowing that $x^* = 0$ which means that $g'(x^*)$ will only consist of terms of g with degree less than 2. And so, we get

$$g'(x^*) = 1 - \left[\frac{2(2-2\lambda) + 0(4\lambda)}{4}\right] = 1 - \frac{4-4\lambda}{4} = \lambda.$$

We want $|g(x^*)| = |\lambda| > 1$. Therefore, when $\lambda < -1$ or $\lambda > 1$, we have a repelling fixed point. Note that we initially conclude that we need $\lambda < 1$ in order for $x = 0$ to be a minimum of $\Phi(x)$ and thus a fixed point of Gauss-Newton method. Thus, only when $\lambda < -1$ is $x = 0$ a repelling fixed point of Gauss-Newton method.

3. For this part, we want to show that if $\lambda < 7/16$, then Levenberg-Marquardt will converge to $x = 0$ i.e. $|g(x = 0)| < 1$.

   Using the $g'(x^*)$ that was computed in the previous part, we can easily note that for Levenberg-Marquardt

   $$g'(x^*) = 1 - \frac{2-2\lambda}{2+\mu}.$$

   We need to show that

   $$\begin{aligned} \left|1 - \frac{2-2\lambda}{2+\mu}\right| &< 1 \\ \Rightarrow 0 \le \frac{1-\lambda}{2+\mu} &\le 1. \end{aligned}$$

   Since $\lambda < \frac{7}{16}$ and $\mu \ge 0$, the fraction will be positive thus we only need to focus on the upper bound.

   $$\begin{aligned} \frac{1-\lambda}{2+\mu} &\le 1 \\ \Rightarrow 2 + \mu &\ge 1 - \lambda \\ \Rightarrow 1 + \mu &\ge -\lambda \end{aligned}$$

   If we further assume that $\mu = |\lambda|$, then this reduces to

   $$\begin{aligned} 1 + \mu &\ge -\lambda \\ \Rightarrow 1 &\ge -\lambda - |\lambda| \end{aligned}$$

   Notice that $-\lambda - |\lambda| \le 0$, therefore the inequality above holds and $|g(x = 0)| < 1$.