Online Lecture Notes

Prof. Boris Houska

April 28, 2022

1 Announcements

The mid-term exam will be about Lecture 1 - 6 + the parts of the "Analysis in Nutshell". This also includes Homework 1 - 6. Tentative plan:

1. Today I will start with Lecture 8 (not relevant for the mid-term exam).
2. Next Tuesday, May 3, we organize a repitition session + Q & A. During this lecture I would go over Lecture 1 - 6 to summarize the most important points:
   (a) Lecture 1: Error Analysis, Numerical and Algorithmic Differentiation
   (b) Lecture 2: Polynomial Interpolation
   (c) Lecture 3: Polynomial Extrapolation, Splines
   (d) Lecture 4: Gauss Approximation
   (e) Lecture 5: Numerical Integration
   (f) "Lecture 6": Analysis in a Nutshell: Vector Spaces, Hilbert Spaces, Cauchy Schwarz inequality, Gram-Schmidt Algorithm
3. Next Thursday, May 5, we would run the mid-term, which will be a take-home exam. The exam will look very similar to mid-term exams from previous years, but you will 24h for solving all problems and sending them back to us. We will start at the beginning of the lecture by distributing the exercise.

Jiahe and Xvting will send the time for the exercise session.

2 Newton's Method and Newton Type Methods

The goal of this lecture is to develop numerical methods for solving the nonlinear equations system

$$f(x) = 0,$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ is a given nonlinear function. The main idea of Newton's method is to recursively approximate f by a linear function at the current iterate x_k,

$$f(x) \approx f(x_k) + M(x_k)(x - x_k)$$

for a suitable invertible Jacobian approximation $M(x) \approx f'(x)$, such that

$$\det(M(x_k)) \neq 0 \; .$$

The case $M(x) = f'(x)$ leads to an exact Newton methods while the case $M(x) \neq f'(x)$ corresponds to a Newton type method. The iteration can be written in the form

$$f(x_k) + M(x_k)(x - x_k) = 0 \qquad \Longleftrightarrow \qquad x_{k+1} = x_k - M(x_k)^{-1} f(x_k) \; .$$

2.1 Scaling Properties of the Exact Newton Method

If we apply the exact Newton method to the scaled equation

$$\tilde{f}(x) = Sf(x) = 0$$

for an invertible matrix S, we obtain the Jacobian $M(x) = \tilde{f}'(x) = Sf'(x)$, which is equivalent to the iteration

$$\begin{aligned} x_{k+1} &= x_k - M(x_k)^{-1}\tilde{f}(x_k) \\ &= x_k - M(x_k)^{-1}Sf(x_k) \\ &= x_k - [Sf'(x_k)]^{-1}Sf(x_k) \\ &= x_k - f'(x_k)^{-1}S^{-1}Sf(x_k) \\ &= x_k - f'(x_k)^{-1}f(x_k). \end{aligned}$$

Thus, Newton's method applied to f and applied to $\tilde{f}$ yield the same iteration! This means that Newton's is invariant under scaling.

2.2 Example

If we have a simple nonlinear function of the form

$$f(x) = e^x - 2,$$

this corresponds to solving the equation

$$0 = f(x) = e^x - 2 \qquad \Longleftrightarrow \qquad x = \ln(2) \; .$$

The corresponding Newton method would be an iteration of the form

$$\begin{aligned} x_{k+1} &= x_k - f'(x_k)^{-1}f(x_k) \\ &= x_k - e^{-x_k}(e^{x_k} - 2) \end{aligned} \tag{1}$$

Numerical experiment: the exact Newton method seems to converge really fast, in just a few steps when, e.g., starting at $x_0 = 1.0$. If we do Newton type iteration, e.g,

$$x_{k+1} \quad = \quad x_k - e^{-0.5}(e^{x_k} - 2), \tag{2}$$

we also observe convergence, but much more iterations are needed. Of course, it is not clear which method is faster for higher dimensional systems, since for large n inverting the exact Jacobian may be expensive.

2.3 Convergence Analysis

In order to analyze the convergence of Newton type methods, we usually make three (arguably all rather debatable) assumptions, namely,

- A1 The equation $f(x^\star) = 0$ has a solution $x^\star \in \mathbb{R}^n$.
- A2 We have an initialization x_0 in a sufficiently small open neighborhood $\mathcal{N}(x^\star) \subseteq \mathbb{R}^n$ of $x^\star$.
- A3 The scaled Jacobians, $M(x_k)^{-1} f'(x)$, are locally Lipschitz continuous with given Lipschitz constant $\omega < \infty$. This means that

$$\forall x, y \in \mathcal{N}(x^\star), \qquad \|M(x_k)^{-1} f'(x) - M(x_k)^{-1} f'(y)\| \leq \omega \|x - y\| \ .$$

The actual convergence rate is obtained by bounding the difference of the iterate of the Newton type method x_k to the solution $x^\star$,

$$x_{k+1} = x_k - M(x_k)^{-1} f(x_k) \ .$$

Thus, we start by working out the norm

$$
\begin{aligned}
\|x_{k+1} - x^\star\| &= \left\|x_k - M(x_k)^{-1} f(x_k) - x^\star\right\| \\
&= \left\|x_k - x^\star - \left[M(x_k)^{-1} f(x_k) - M(x_k)^{-1} f(x^\star)\right]\right\| \\
&= \left\|x_k - x^\star - M(x_k)^{-1} \int_0^1 f'(x^\star + s(x_k - x^\star))(x_k - x^\star)\,\mathrm{d}s\right\| \\
&= \left\|x_k - x^\star - M(x_k)^{-1} \int_0^1 f'(x_k)(x_k - x^\star)\,\mathrm{d}s\right. \\
&\quad \left. + \int_0^1 M(x_k)^{-1}\left[f'(x_k) - f'(x^\star + s(x_k - x^\star))\right](x_k - x^\star)\,\mathrm{d}s\right\| \\
&\le \left\|x_k - x^\star - M(x_k)^{-1} \int_0^1 f'(x_k)(x_k - x^\star)\,\mathrm{d}s\right\| \\
&\quad + \left\|\int_0^1 M(x_k)^{-1}\left[f'(x_k) - f'(x^\star + s(x_k - x^\star))\right](x_k - x^\star)\,\mathrm{d}s\right\| \\
&= \left\|x_k - x^\star - M(x_k)^{-1} f'(x_k)(x_k - x^\star)\right\| \\
&\quad + \left\|\int_0^1 M(x_k)^{-1}\left[f'(x_k) - f'(x^\star + s(x_k - x^\star))\right](x_k - x^\star)\,\mathrm{d}s\right\| \\
&\le \left\|x_k - x^\star - M(x_k)^{-1} f'(x_k)(x_k - x^\star)\right\| \\
&\quad + \int_0^1 \left\|M(x_k)^{-1}\left[f'(x_k) - f'(x^\star + s(x_k - x^\star))\right](x_k - x^\star)\right\|\,\mathrm{d}s \\
&\overset{\text{A3}}{\le} \left\|x_k - x^\star - M(x_k)^{-1} f'(x_k)(x_k - x^\star)\right\| + \int_0^1 s\omega\,\|(x_k - x^\star)\|^2\,\mathrm{d}s \\
&= \left\|x_k - x^\star - M(x_k)^{-1} f'(x_k)(x_k - x^\star)\right\| + \frac{\omega}{2}\,\|(x_k - x^\star)\|^2 \\
&= \left\|M(x_k)^{-1}\left[M(x_k) - f'(x_k)\right](x_k - x^\star)\right\| + \frac{\omega}{2}\,\|(x_k - x^\star)\|^2 \\
&\le \kappa\|x_k - x^\star\| + \frac{\omega}{2}\,\|(x_k - x^\star)\|^2 \qquad (3)
\end{aligned}
$$

with κ being a bound on the scaled Jacobian approximation

$$\left\|M(x_k)^{-1}\left[M(x_k) - f'(x_k)\right]\right\| \;\le\; \kappa$$

In summary, if the Jacobian approximation is sufficiently accurate, such that $\kappa < 1$, the Newton method is locally convergent, since

$$\|x_{k+1} - x^\star\| \;\le\; \kappa\|x_k - x^\star\| + \frac{\omega}{2}\,\|(x_k - x^\star)\|^2 \;.$$

Thus, in general, Newton type methods converge linearly, depending on κ. But the exact Newton methods is obtained for $\kappa = 0$. In this case, we get quadratic convergence.