

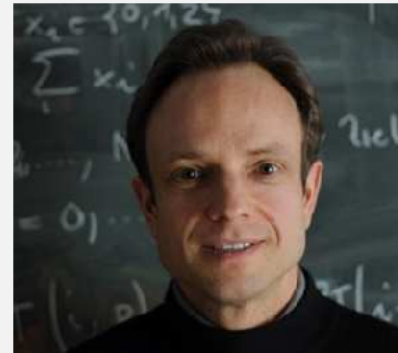
Course Video

Section 8.7.2 , Section 8.7.3: (Trust-region)
Gauss-Newton Method

Prof. R. Hiptmair, SAM, ETH Zurich

Date: December 6, 2020

(C) Seminar für Angewandte Mathematik, ETH Zürich



Dependencies. Section 3.3.4

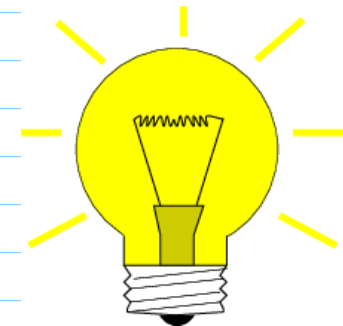
Duration: 22 minutes

VIII Iterative Methods for Non-Linear Systems of Equations

8.7 Non-Linear Least Squares

Non-linear least squares problem $\Leftrightarrow \mathcal{R}(\underline{x}) = 0$ Given: $F : D \subset \mathbb{R}^n \mapsto \mathbb{R}^m$, $m, n \in \mathbb{N}$, $m > n$.Find: $\mathbf{x}^* \in D$: $\mathbf{x}^* = \operatorname{argmin}_{\mathbf{x} \in D} \Phi(\mathbf{x})$, $\Phi(\mathbf{x}) := \frac{1}{2} \|F(\mathbf{x})\|_2^2$. (8.7.0.7)

8.7.2. Gauss-Newton Method



Idea:

Local linearization of F : $F(\mathbf{x}) \approx F(\mathbf{y}) + DF(\mathbf{y})(\mathbf{x} - \mathbf{y})$

➤ Leads to a sequence of linear least squares problems

↓
iterative method

(2)

$\operatorname{argmin}_{\mathbf{x} \in \mathbb{R}^n} \|F(\mathbf{x})\|_2$ is approximated by $\underbrace{\operatorname{argmin}_{\mathbf{x} \in \mathbb{R}^n} \|F(\mathbf{x}_0) + DF(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\|_2}_{(\spadesuit)}$

$(\spadesuit) \Leftrightarrow \operatorname{argmin}_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|$ with $\mathbf{A} := DF(\mathbf{x}_0) \in \mathbb{R}^{m,n}$, [Jacobian]
 $\mathbf{b} := -F(\mathbf{x}_0) + DF(\mathbf{x}_0)\mathbf{x}_0 \in \mathbb{R}^m$.

linear least-squares problem

Gauss-Newton iteration (under full-rank assumption (8.7.0.9))

Initial guess $\mathbf{x}^{(0)} \in D$

$$\mathbf{x}^{(k+1)} := \mathbf{x}^{(k)} - \mathbf{s}, \quad \mathbf{s} := \operatorname{argmin}_{\mathbf{x} \in \mathbb{R}^n} \|F(\mathbf{x}^{(k)}) - DF(\mathbf{x}^{(k)})\mathbf{x}\|_2. \quad (8.7.2.1)$$

Gauss-Newton correction $\rightarrow$ a linear least squares problem!

C++ code 8.7.2.2: Template for Gauss-Newton method

```

2 template <class Function, class Jacobian>
3 VectorXd gn(const Eigen::VectorXd &init, Function &&F, Jacobian &&J,
4             double rtol = 1.0E-6, double atol = 1.0E-8) {
5     Eigen::VectorXd x = init; // Vector for iterates  $\mathbf{x}^{(k)}$ 
6     // Vector for Gauss-Newton correction  $\mathbf{s}$ 
7     Eigen::VectorXd s = J(x).householderQr().solve(F(x)); //
8     x = x - s;
9     // A posteriori termination based on absolute and relative tolerances
10    while ((s.norm() > rtol * x.norm()) && (s.norm() > atol)) { //
11        s = J(x).householderQr().solve(F(x)); //  $\rightarrow$  linear LSQ
12        x = x - s;
13    }
14    return x;
15 }
```

correction based

Rem. 8.7.2.3 (Gauss-Newton vs. Newton)

+ 2nd-derivative free
 + Larger domain of conv.

+ local quadratic conv.

Ex. 8.7.2.4 (Non-linear fitting of data)

Fitting functions $f(x_1, x_2, x_3, t) = x_1 + x_2 e^{-x_3 t}$
 $\triangleright$ Overdetermined NLSIE $F(\mathbf{x}) = 0$

$$F: \mathbb{R}^3 \mapsto \mathbb{R}^m, \quad F(\mathbf{x}) := \begin{bmatrix} x_1 + x_2 \exp(-x_3 t_1) - y_1 \\ \vdots \\ x_1 + x_2 \exp(-x_3 t_m) - y_m \end{bmatrix}.$$

21 data points,

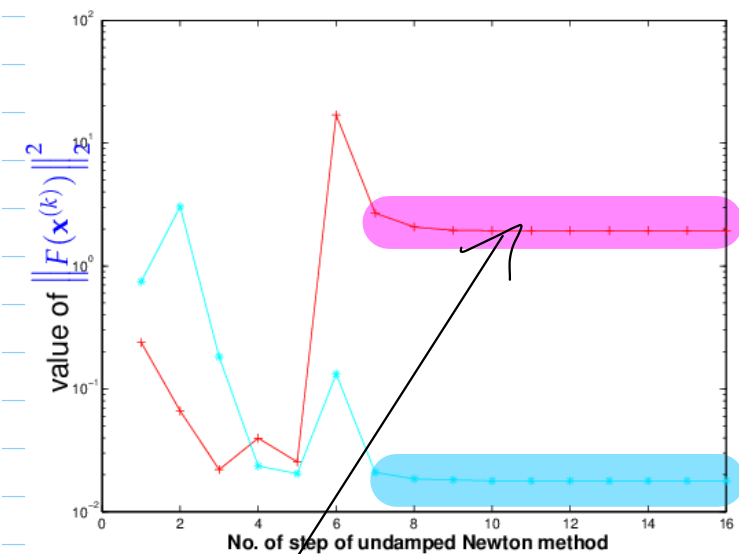
"random"

♦ initial value $(1.8, 1.8, 0.1)^T$ (red curves, blue curves)

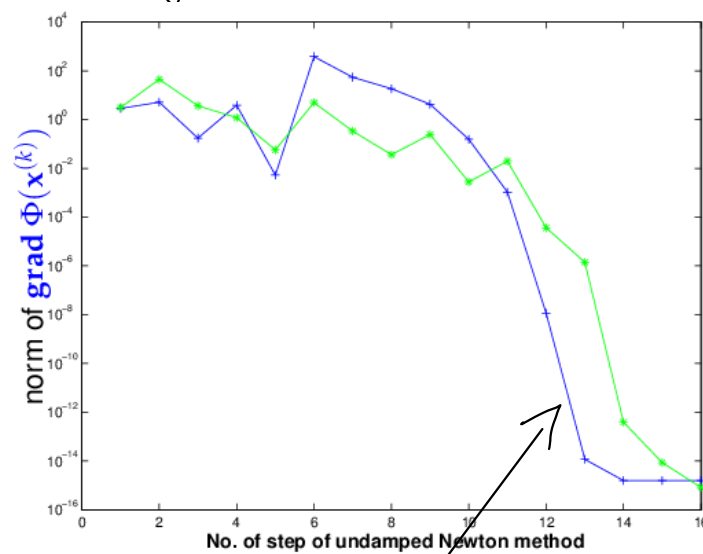
♦ initial value $(1.5, 1.5, 0.1)^T$ (cyan curves, green curves)

③

- Newton method for $\text{grad } \phi(x) = 0$

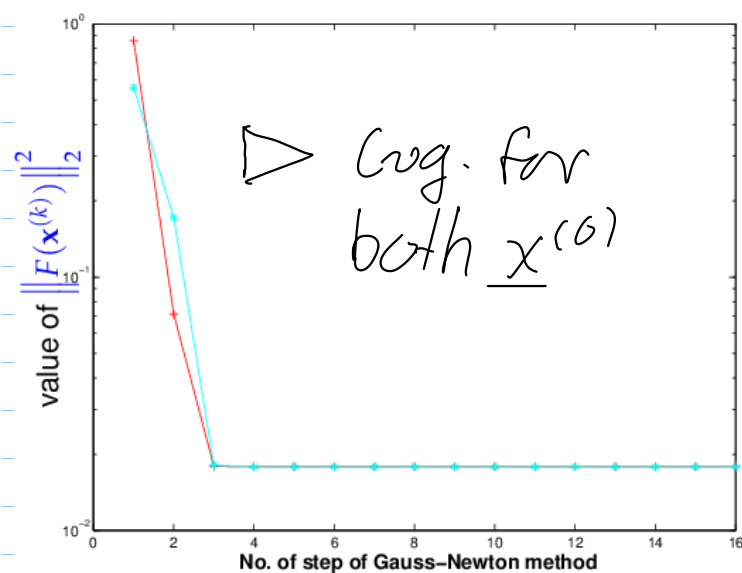


Newton it. stuck in local min.

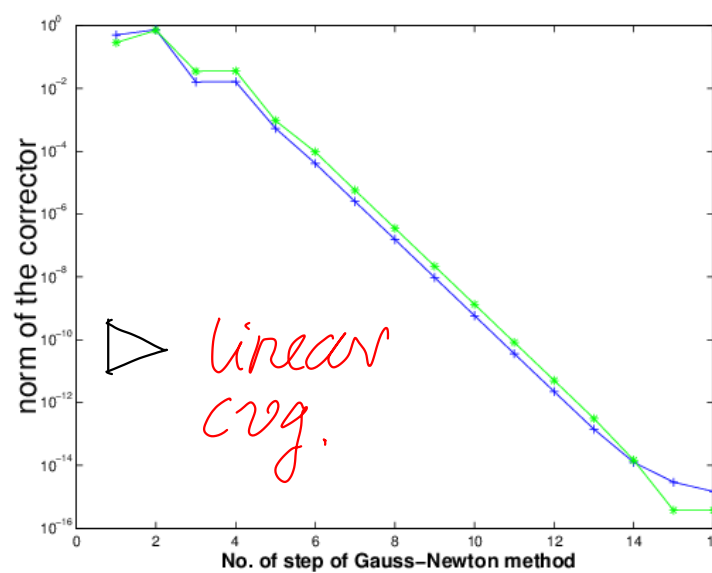


quadratic conv.

- Gauss-Newton method



▷ Cvg. for both $\underline{x}^{(0)}$



▷ linear conv.

8.7.3. Trust Region Method

$\hat{=}$ Implicitly damped Gauss-Newton method

Idea: damping of the Gauss-Newton correction in (8.7.2.1) using a **penalty term**

instead of $\|F(\mathbf{x}^{(k)}) + DF(\mathbf{x}^{(k)})\mathbf{s}\|_2^2$ minimize $\|F(\mathbf{x}^{(k)}) + DF(\mathbf{x}^{(k)})\mathbf{s}\|_2^2 + \lambda \|\mathbf{s}\|_2^2$,

where $\lambda > 0$ is a **penalty parameter**.

$$\begin{cases} \underline{x}^{(k+1)} = \underline{x}^{(k)} - \underline{s} \\ \mathbf{s} := \underset{\mathbf{z} \in \mathbb{R}^n}{\text{argmin}} \|F(\mathbf{x}^{(k)}) + DF(\mathbf{x}^{(k)})\mathbf{z}\|_2^2 + \lambda \|\mathbf{z}\|_2^2 \end{cases}$$

↳ Extended LSG problem

Normal equations:

$$(DF(\mathbf{x}^{(k)})^T DF(\mathbf{x}^{(k)}) + \lambda \mathbf{I}) \mathbf{s} = -DF(\mathbf{x}^{(k)}) F(\mathbf{x}^{(k)}) \quad (8.7.3.2)$$

$\lambda \approx ?$ Heuristics, $\lambda \rightarrow 0$ as $\underline{x}^{(k)} \rightarrow \underline{x}^*$

④

▷ Review questions 8.7.3.3.

5

7