

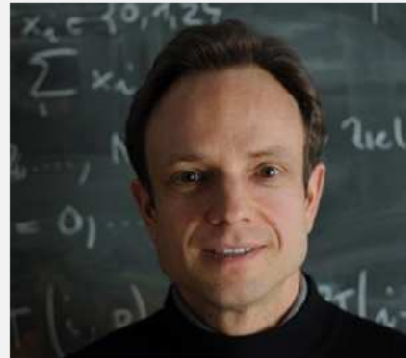
Course Video

Section 8.6: Quasi-Newton Method

Prof. R. Hiptmair, SAM, ETH Zurich

Date: December 6, 2020

(C) Seminar für Angewandte Mathematik, ETH Zürich



Dependencies. Section 8.5.1, Section 2.6

Duration: 27 minutes



Updated version compared to video

VIII Iterative Methods for Non-Linear Systems of Equations

8.6 Quasi-Newton Method

$$\text{NLSE: } \underline{x}^* \in D : F(\underline{x}^*) = 0$$

$$F : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

Drawback of Newton method : DF required

Goal : Derivative-free iteration of NLSE

Achieved for $n = 1$: Secant method, 8.4.2.3

$$x^{(k+1)} = x^{(k)} - \frac{x^{(k)} - x^{(k-1)}}{F(x^{(k)}) - F(x^{(k-1)})} F(x^{(k)})$$

$$\text{N.I: } x^{(k+1)} = x^{(k)} - \frac{1}{F'(x^{(k)})} F(x^{(k)})$$

(2)

 $n = 1$:

$$DF(x^{(k)}) \longrightarrow \frac{F(x^{(k)}) - F(x^{(k-1)})}{x^{(k)} - x^{(k-1)}} =: J_k$$

 $n > 1$:

Replace $DF(x^{(k)})$ w/ "different quotient approximation" $J_k \in \mathbb{R}^{n,n}$

Idea: rewrite (8.6.0.1) as a **secant condition** for an approximation $J_k \approx DF(x^{(k)})$, $x^{(k)} \triangleq$ the current iterate:

(SC)

$$J_k(x^{(k)} - x^{(k-1)}) = F(x^{(k)}) - F(x^{(k-1)}) \quad (8.6.0.2)$$

$$\text{Iteration: } x^{(k+1)} := x^{(k)} - J_k^{-1} F(x^{(k)}) \quad (8.6.0.3)$$

does not fix J_k : Reuse J_{k-1} : $J_k \equiv J_{k-1}$ on "non-secant directions"

► Broyden's conditions: $J_k z = J_{k-1} z \quad \forall z: z \perp (x^{(k)} - x^{(k-1)})$. (BC)

$$J_k := J_{k-1} + \frac{F(x^{(k)})(x^{(k)} - x^{(k-1)})^\top}{\|x^{(k)} - x^{(k-1)}\|_2^2} \quad (8.6.0.5)$$

(BC) + (SC) unique determine J_k

Final form of **Broyden's quasi-Newton method** for solving $F(x) = 0$:

$$x^{(k+1)} := x^{(k)} + \Delta x^{(k)}, \quad \Delta x^{(k)} := -J_k^{-1} F(x^{(k)}),$$

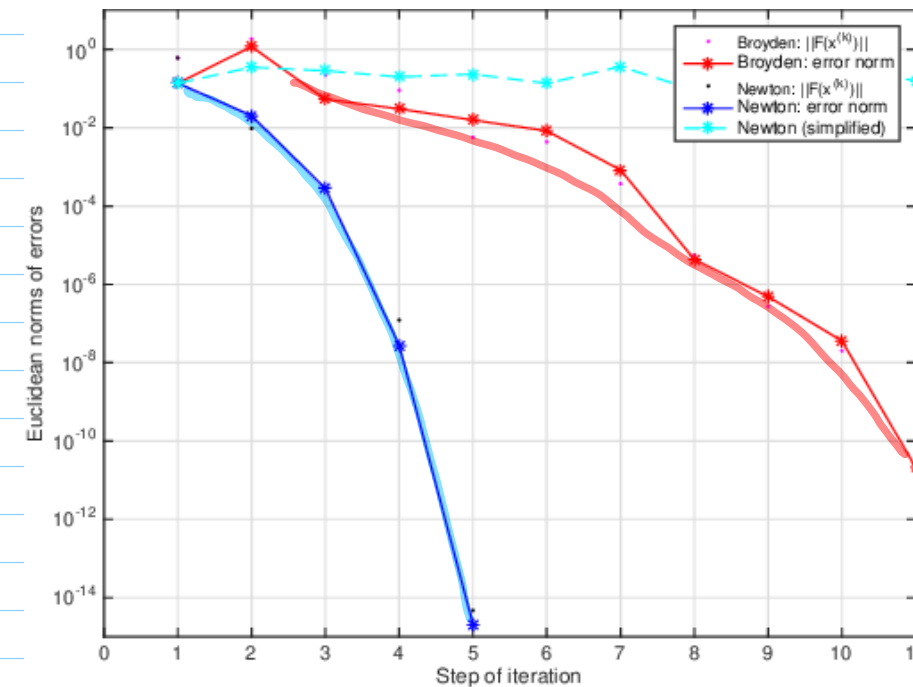
$$J_{k+1} := J_k + \frac{F(x^{(k+1)})(\Delta x^{(k)})^\top}{\|\Delta x^{(k)}\|_2^2} \approx DF(x^{(k)}) \quad (8.6.0.6)$$

req.: initial guess J_0 : $J_0 := DF(x^{(0)})$

Exp. 8.6.0.10 (Convergence of Broyden's method)

$$F(x) = \begin{bmatrix} x_1^2 - x_2^4 \\ x_1 - x_2^3 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \quad \text{with solution } F\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = 0, \quad (8.5.2.2)$$

and take $x^{(0)} = [0.7, 0.7]^T$. As starting value for the matrix iteration we use $J_0 = DF(x^{(0)})$.



◁ Lin-log error plot

▷ No asymptotic quadratic conv.

③

§ 8.6.0.15 (Implementation of Broyden's quasi-Newton method)

Final form of **Broyden's quasi-Newton method** for solving $F(\mathbf{x}) = 0$:

$$\begin{aligned} \mathbf{x}^{(k+1)} &:= \mathbf{x}^{(k)} + \Delta \mathbf{x}^{(k)}, \quad \Delta \mathbf{x}^{(k)} := -\mathbf{J}_k^{-1} F(\mathbf{x}^{(k)}), \\ \mathbf{J}_{k+1} &:= \mathbf{J}_k + \frac{F(\mathbf{x}^{(k+1)}) (\Delta \mathbf{x}^{(k)})^\top}{\|\Delta \mathbf{x}^{(k)}\|_2^2} \stackrel{!}{=} \text{rank-1-modification} \quad (8.6.0.6) \end{aligned}$$

Lemma 2.6.0.21. Sherman-Morrison-Woodbury formula

For regular $\mathbf{A} \in \mathbb{K}^{n,n}$, and $\mathbf{U}, \mathbf{V} \in \mathbb{K}^{n,k}$, $n, k \in \mathbb{N}$, $k \leq n$, holds

$$(\mathbf{A} + \mathbf{U}\mathbf{V}^H)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{I} + \mathbf{V}^H\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}^H\mathbf{A}^{-1},$$

if $\mathbf{I} + \mathbf{V}^H\mathbf{A}^{-1}\mathbf{U}$ is regular.

Use for $k=1$:

$$(\mathbf{A} + \mathbf{u}\mathbf{v}^\top)^{-1} = \left(\mathbf{I} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^\top}{1 + \mathbf{v}^\top\mathbf{A}^{-1}\mathbf{u}} \right) \mathbf{A}^{-1}, \quad 1 + \mathbf{v}^\top\mathbf{A}^{-1}\mathbf{u} \neq 0,$$

with $\mathbf{u} := F(\mathbf{x}^{(k+1)})$, $\mathbf{v} := \frac{\Delta \mathbf{x}^{(k)}}{\|\Delta \mathbf{x}^{(k)}\|_2^2}$

▷ Update formula for $\mathbf{J}_k^{-1}$

$$\mathbf{J}_{k+1}^{-1} = \left(\mathbf{I} - \frac{\mathbf{J}_k^{-1} F(\mathbf{x}^{(k+1)}) (\Delta \mathbf{x}^{(k)})^\top}{\|\Delta \mathbf{x}^{(k)}\|_2^2 + (\Delta \mathbf{x}^{(k)})^\top \mathbf{J}_k^{-1} F(\mathbf{x}^{(k+1)})} \right) \mathbf{J}_k^{-1} = \left(\mathbf{I} - \frac{\Delta \bar{\mathbf{x}}^{(k+1)} (\Delta \mathbf{x}^{(k)})^\top}{\|\Delta \mathbf{x}^{(k)}\|_2^2 + (\Delta \mathbf{x}^{(k)})^\top \Delta \bar{\mathbf{x}}^{(k+1)}} \right) \mathbf{J}_k^{-1}, \quad (8.6.0.16)$$

↑ simplified quasi-Newton correction
(→ useful for termination & damping)

Req.: $|\Delta \mathbf{x}^{(k)\top} \Delta \bar{\mathbf{x}}^{(k+1)}| < \|\Delta \mathbf{x}^{(k)}\|_2^2$

$$\Delta \mathbf{x}^{(k)} = -\mathbf{J}_k^{-1} F(\mathbf{x}^{(k)}) = -\prod_{\ell=0}^{k-1} \left(\mathbf{I} - \frac{\Delta \bar{\mathbf{x}}^{(\ell+1)} (\Delta \mathbf{x}^{(\ell)})^\top}{\|\Delta \mathbf{x}^{(\ell)}\|_2^2 + (\Delta \mathbf{x}^{(\ell)})^\top \Delta \bar{\mathbf{x}}^{(\ell+1)}} \right) \mathbf{J}_0^{-1} F(\mathbf{x}^{(k)}).$$

→ LM-factorize once: $O(n^3)$

$\mathbf{t}_0 \in \mathbb{R}^n$: $\mathbf{J}_0 \mathbf{t}_0 = F(\mathbf{x}^{(k)})$, → 2 triangular solves $O(n^2)$

$$\mathbf{t}_{\ell+1} := \mathbf{t}_\ell - \Delta \bar{\mathbf{x}}^{(\ell+1)} \frac{(\Delta \mathbf{x}^{(\ell)})^\top \mathbf{t}_\ell}{\|\mathbf{x}^{(\ell)}\|_2^2 + (\Delta \mathbf{x}^{(\ell)})^\top \Delta \bar{\mathbf{x}}^{(\ell+1)}}, \quad \ell = 0, \dots, k-1, \quad (8.6.0.18)$$

→ $O(kn)$

$\Delta \mathbf{x}^{(k)} := -\mathbf{t}_k$,

Total cost = $O(n^3 + Nn^2 + N^2n)$
(N steps)

(4)

▷ Review questions 8.6.0.22

5

7