Levenberg-Marquardt-Verfahren

Probleme Gauß-Newton

- Fälle in denen $F'(x)$ nicht vollen Rang hat
- Überschießen möglich

Lösung: Levenberg-Marquardt-Verfahren

- Erweiterung Matrix $\rightarrow$ stets voller Rang $\rightarrow$ linearisiertes Problem stets eindeutig lösbar
- Dämpfungsstrategie

Ansatz und Algorithmus siehe Folien 6.15 und 6.17.

Bem: Konvergenzordnung (normalerweise) $p = 1$

Beispiel: Levenberg-Marquardt-Verfahren

Gegeben:

- Modell:

$$(\hat{x} - a)^2 + \mathrm{e}^{b\left(\hat{x}^2 + \hat{y}^2\right)} - 5 = 0$$

- Messwerte:

$\hat{x}_i$	2	3	4
$\hat{y}_i$	0	2	0

Gesucht:

Parameter $x = \begin{pmatrix} a \\ b \end{pmatrix}$

Vorgehen:

Bestimme $F(x^k)$, $F'(x^k)$ durch Einsetzen der Messwerte in das Modell:

$$F(x^k) = \begin{bmatrix} (2-a)^2 + e^{b(2^2+0^2)} - 5 \\ (3-a)^2 + e^{b(3^2+2^2)} - 5 \\ (4-a)^2 + e^{b(4^2+0^2)} - 5 \end{bmatrix}$$

$$F'(x^k) = \begin{bmatrix} -4+2a & 4e^{4b} \\ -6+2a & 13e^{13b} \\ -8+2a & 16e^{16b} \end{bmatrix}$$

Hinweis: In der Praxis müssen $F(x^k)$ und $F'(x^k)$ nicht in dieser ausführlichen Form aufgestellt werden. Es genügt jeweils eine Zeile $F_i(x^k)$ und $F_i'(x^k)$ zu kennen. Sämtliche Werte für a, b, $\hat{x}_i$ und $\hat{y}_i$ werden dann während des Programmablaufs eingesetzt.

Löse in jedem Schritt das lineare Ausgleichsproblem:

$$\left\| \begin{pmatrix} F'(x^k) \\ \mu I \end{pmatrix} s^k + \begin{pmatrix} F(x^k) \\ 0 \end{pmatrix} \right\|_2 \to min$$

Im Anschluss daran wird ρ_μ für den aktuellen Schritt k berechnet:

$$\rho_\mu := \frac{\|F(x^k)\|_2^2 - \|F(x^k+s^k)\|_2^2}{\|F(x^k)\|_2^2 - \|F(x^k) + F'(x^k)s^k\|_2^2} =: \frac{\triangle R\left(x^k, s^k\right)}{\triangle \tilde{R}\left(x^k, s^k\right)}$$

Hierin bedeuten

- $\triangle R\left(x^k, s^k\right):$ *Änderung tatsächliches Residuum*
- $\triangle \tilde{R}\left(x^k, s^k\right):$ *Änderung Residuum lineares Modell* $\left(\triangle \tilde{R}\left(x^k, s^k\right) \geq 0\right)$

Akzeptable Korrektur: $\triangle R\left(x^k, s^k\right) > 0$, d.h. Fehler wird kleiner.

Abhängig von ρ_μ wird im Anschluss entschieden, ob die berechnete Korrektur s^k beibehalten oder verworfen wird und inwiefern der Dämpfungsparameter μ angepasst werden muss. Es gilt $\beta_0, \beta_1 \in \mathbb{R}$ mit $0 < \beta_0 < \beta_1 < 1$. Hier verwenden wir $\beta_0 = 0.2$ und $\beta_1 = 0.8$.

- $\rho_\mu \leq \beta_0$: s^k wird nicht akzeptiert; μ wird verdoppelt (im Allgemeinen auch andere Wahl möglich) und es wird eine neue Korrektur s^k berechnet.
- $\beta_0 < \rho_\mu < \beta_1$: s^k wird akzeptiert; μ wird beibehalten
- $\rho_\mu \geq \beta_1$: s^k wird akzeptiert (nicht neu berechnen); μ wird halbiert (im Allgemeinen auch andere Wahl möglich)

Falls s^k verworfen wird (Fall: $\rho_\mu \leq \beta_0$), wird das lineare Ausgleichsproblem mit dem neuen Wert für μ erneut aufgestellt und s^k neu berechnet.

Durchführung:

Startwert:

$$x^0 = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$$

Gestartet wird mit $\mu^0 = 1$. Der erste Schritt wird im Folgenden ausführlich beschrieben:

- Zunächst werden $F'(x^0)$ und $F(x^0)$ aufgestellt:

$$F'(x^0) = \begin{bmatrix} -4 + 2 \cdot 4 & 4e^{4 \cdot 0} \\ -6 + 2 \cdot 4 & 13e^{13 \cdot 0} \\ -8 + 2 \cdot 4 & 16e^{16 \cdot 0} \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 2 & 13 \\ 0 & 16 \end{bmatrix}$$

$$F(x^0) = \begin{bmatrix} (2-4) + e^{0 \cdot (2^2+0^2)} - 5 \\ (3-4)^2 + e^{0 \cdot (3^2+2^2)} - 5 \\ (4-4)^2 + e^{0 \cdot (4^2+0^2)} - 5 \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \\ -4 \end{bmatrix}$$

- Anschließend wird das lineare Ausgleichsproblem aufgestellt:

$$\left\| \begin{pmatrix} F'(x^0) \\ \mu I \end{pmatrix} s^0 + \begin{pmatrix} F(x^0) \\ 0 \end{pmatrix} \right\|_2 = \left\| \begin{pmatrix} 4 & 4 \\ 2 & 13 \\ 0 & 16 \\ \mu & 0 \\ 0 & \mu \end{pmatrix} s^0 + \begin{pmatrix} 0 \\ -3 \\ -4 \\ 0 \\ 0 \end{pmatrix} \right\|_2 \to min$$

- Mit dem Dämpfungsparameter $\mu^0 = 1$ erhält man als Lösung des linearen Ausgleichsproblems

$$s^0 = \begin{pmatrix} -0.22266... \\ 0.25418... \end{pmatrix}$$

- Mit dem (vorläufigen) Wert für $x^1 = x^0 + s^0$ lässt sich $F(x^1) = F(x^0 + s^0)$ bestimmen:

$$x^1 = \begin{pmatrix} 4 \\ 0 \end{pmatrix} + \begin{pmatrix} -0.22266... \\ 0.25418... \end{pmatrix} = \begin{pmatrix} 3.777334398 \\ 0.2541899441 \end{pmatrix}$$

$$F(x^1) = \begin{bmatrix} (2-3.777...) + e^{0.2541... \cdot (2^2+0^2)} - 5 \\ (3-3.777...)^2 + e^{0.2541... \cdot (3^2+2^2)} - 5 \\ (4-3.777...)^2 + e^{0.2541... \cdot (4^2+0^2)} - 5 \end{bmatrix} = \begin{bmatrix} 0.92314... \\ 22.838... \\ 53.433... \end{bmatrix}$$

- Zudem muss $F(x^0) + F'(x^0)s^0$ bestimmt werden:

$$\begin{aligned} F(x^0) + F'(x^0)s^0 &= \begin{bmatrix} (2-4) + e^{0 \cdot (2^2+0^2)} - 5 \\ (3-4)^2 + e^{0 \cdot (3^2+2^2)} - 5 \\ (4-4)^2 + e^{0 \cdot (4^2+0^2)} - 5 \end{bmatrix} + \begin{bmatrix} -4 + 2 \cdot 4 & 4e^{4 \cdot 0} \\ -6 + 2 \cdot 4 & 13e^{13 \cdot 0} \\ -8 + 2 \cdot 4 & 16e^{16 \cdot 0} \end{bmatrix} \begin{pmatrix} -0.22266... \\ 0.25418... \end{pmatrix} \\ &= \begin{pmatrix} 0.12609... \\ -0.14086... \\ 0.067039... \end{pmatrix} \end{aligned}$$

- Mit diesen Werten kann ρ_{μ_0} berechnet werden:

$$\rho_{\mu_0} := \frac{\|F(x^0)\|_2^2 - \|F(x^0 + s^0)\|_2^2}{\|F(x^0)\|_2^2 - \|F(x^0) + F'(x^0)s^0\|_2^2} = -134.3190...$$

- Da $\rho_\mu = -134.3190... \leq \beta_0 = 0.2$ werden s^0 und $x^1 = x^0 + s^0$ verworfen. Der Dämpfungsparameter μ wird verdoppelt, d.h. $\mu^0 = 2\mu^0 = 2 \cdot 1$, und das lineare Ausgleichsproblem wird erneut aufgestellt. Anschließend wird ρ_{μ_0} erneut berechnet und entschieden, ob die neue Lösung s^0 verwendet wird. Dieser Vorgang wird ggfs. so lange wiederholt, bis der Fall $\beta_0 < \rho_\mu < \beta_1$ oder $\rho_\mu \geq \beta_1$ eintritt. Der in diesen Fällen berechnete Wert $x^1 = x^0 + s^0$ wird beibehalten und für den neuen Iterationsschritt $k = 1$ verwendet.
- Im Folgenden sind die weiteren Iterationsschritte tabellarisch aufgelistet. Zu beachten ist, dass der Fall $\rho_\mu \leq \beta_0$ im betrachteten Beispiel nur in der Iteration $k = 0$ auftritt. Daher muss die Lösung x^k nur in dieser Iteration (mehrmals) verworfen werden. Ab Iteration 7 liegen im Zähler und Nenner von ρ_μ sehr kleine Werte vor, weshalb Auslöschung auftritt. In Iteration 8 und 9 wird ρ_μ asymptotisch zu 1.0 gesetzt.

Iteration k	ρ_{μ_k}	**neues** μ^{k+1}	x^{k+1} **(ggfs. verworfene Werte)**
0	-134.3190547	2	$\begin{pmatrix} 3.777334398 \\ 0.2541899441 \end{pmatrix}$ (verworfen)
	-112.3409631	4	$\begin{pmatrix} 3.814266488 \\ 0.2489905787 \end{pmatrix}$ (verworfen)
	-69.95301287	8	$\begin{pmatrix} 3.892156864 \\ 0.2352941176 \end{pmatrix}$ (verworfen)
	-24.48620988	16	$\begin{pmatrix} 3.968122786 \\ 0.2066115702 \end{pmatrix}$ (verworfen)
	-0.7462026156	32	$\begin{pmatrix} 3.999244523 \\ 0.1478217073 \end{pmatrix}$ (verworfen)
	1.537651109	16	$\begin{pmatrix} 4.002922047 \\ 0.07022339523 \end{pmatrix}$
1	0.9410590343	8	$\begin{pmatrix} 3.997152462 \\ 0.1080604032 \end{pmatrix}$
2	0.9969713628	4	$\begin{pmatrix} 3.979022175 \\ 0.1024608243 \end{pmatrix}$
3	0.9942238019	2	$\begin{pmatrix} 3.945533698 \\ 0.1025966463 \end{pmatrix}$
4	0.9962250017	1	$\begin{pmatrix} 3.920827151 \\ 0.1028660524 \end{pmatrix}$
5	0.9973927375	0.5	$\begin{pmatrix} 3.915354567 \\ 0.1029146520 \end{pmatrix}$
6	0.9970614693	0.25	$\begin{pmatrix} 3.915046211 \\ 0.1029172713 \end{pmatrix}$
7	2.000000000	0.125	$\begin{pmatrix} 3.915042531 \\ 0.1029172980 \end{pmatrix}$
8	1.0	0.0625	$\begin{pmatrix} 3.915042527 \\ 0.1029172979 \end{pmatrix}$
9	1.0	0.03125	$\begin{pmatrix} 3.915042527 \\ 0.1029172979 \end{pmatrix}$

Bezogen auf die gewählte Anzahl an signifikanten Stellen erfolgte von der 8. zur 9. Iteration keine Veränderung mehr. Man erhält das Ergebnis: $(\hat{x} - 3.915042527)^2 + \mathrm{e}^{0.1029172979(\hat{x}^2+\hat{y}^2)} - 5 = 0$