

# Veranstaltung Titel

Vortragender

Assistenten

Institut XXX

Wintersemester/Sommersemester xyz

# Übersicht

Themen: Dahmen & Reusken Kap 6.1-6.2

- ▶ Das nichtlineare Ausgleichsproblem
- ▶ Gauß-Newton-Verfahren

# Übersicht

Themen: Dahmen & Reusken Kap 6.1-6.2

- ▶ Das nichtlineare Ausgleichsproblem
- ▶ Gauß-Newton-Verfahren

Was Sie mitnehmen sollten:

- ▶ Wie ist die Problemstellung bei einem nichtlinearen Ausgleichsproblem
- ▶ Wie funktioniert das Gauß-Newton-Verfahren
- ▶ Wichtige (Konvergenz-)Eigenschaften dieser Methode

# Problemstellung

Vorgegebenes Modell

$$b(t) = y(t; x_1, \dots, x_n)$$

mit **unbekannten Parameterwerten**  $x_1, \dots, x_n$ .

Verfügbare Meßdaten:  $b_i \approx b(t_i)$ ,  $i = 1, \dots, m$  ( $m > n$ ).

# Problemstellung

Vorgegebenes Modell

$$b(t) = y(t; x_1, \dots, x_n)$$

mit **unbekannten Parameterwerten**  $x_1, \dots, x_n$ .

Verfügbare Meßdaten:  $b_i \approx b(t_i)$ ,  $i = 1, \dots, m$  ( $m > n$ ).

## Gauß'sche Fehlerquadratmethode

Bestimme diejenige Parameter  $x_1^*, \dots, x_n^*$ , die das Residuum in der Euklidischen Norm minimieren:

$$\sum_{i=1}^m (y(t_i; x_1^*, \dots, x_n^*) - b_i)^2 = \min_{x \in \mathbb{R}^n} \sum_{i=1}^m (y(t_i; x_1, \dots, x_n) - b_i)^2$$

# Problemstellung

Vorgegebenes Modell

$$b(t) = y(t; x_1, \dots, x_n)$$

mit **unbekannten Parameterwerten**  $x_1, \dots, x_n$ .

Verfügbare Meßdaten:  $b_i \approx b(t_i)$ ,  $i = 1, \dots, m$  ( $m > n$ ).

## Gauß'sche Fehlerquadratmethode

Bestimme diejenige Parameter  $x_1^*, \dots, x_n^*$ , die das Residuum in der Euklidischen Norm minimieren:

$$\sum_{i=1}^m (y(t_i; x_1^*, \dots, x_n^*) - b_i)^2 = \min_{x \in \mathbb{R}^n} \sum_{i=1}^m (y(t_i; x_1, \dots, x_n) - b_i)^2$$

Falls die Parameter linear in  $y$  eingehen: **lineare Ausgleichsrechnung**.  
Nichtlineare Abhängigkeiten: **nichtlineares Ausgleichsproblem**.

# Zur Erinnerung: Lineares Ausgleichsproblem

Im linearen Fall

$$y(t_i; x_1, \dots, x_n) = a_{i,1}x_1 + \dots + a_{i,n}x_n$$

Ein lineares Ausgleichsproblem in Matrix-Vektor Notation:

$$\|A x^* - b\|_2 = \min_{x \in \mathbb{R}^n} \|A x - b\|_2$$

# Zur Erinnerung: Lineares Ausgleichsproblem

Im linearen Fall

$$y(t_i; x_1, \dots, x_n) = a_{i,1}x_1 + \dots + a_{i,n}x_n$$

Ein lineares Ausgleichsproblem in Matrix-Vektor Notation:

$$\|A x^* - b\|_2 = \min_{x \in \mathbb{R}^n} \|A x - b\|_2$$

Fallunterscheidung:

- ▶ **Rang**( $A$ ) =  $n$  (gewöhnliches lineares Ausgleichsproblem):  
obige Aufgabe hat eine eindeutige Lösung  $x^*$ .
- ▶ Keine Voraussetzung an  $A$  (allgemeines lineares Ausgleichsproblem):  
obige Aufgabe hat eine eindeutige Lösung  $x^*$  mit minimaler Euklidischer Norm.

# Beispiel 6.1

Differentialgleichung einer gedämpften Schwingung:

$$m u'' + b u' + D u = 0,$$

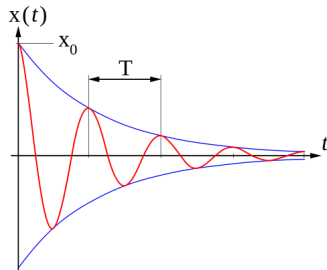
mit Masse  $m$ , Dämpfungskonstante  $b$  und Federkonstante  $D$ .

Lösungen haben die **nichtlineare** Form:

$$u(t) = u_0 e^{-\delta t} \sin(\omega_d t + \varphi_0),$$

wobei:

$u_0$	...	Anfangswert
$\varphi_0$	...	Nullphasenwinkel
$\delta$	...	Abklingkonstante
$\omega_d$	...	Eigenkreisfrequenz



# Beispiel 6.1

## Gegeben:

- ▶ 10 Messungen an den Punkten  $t_1, t_2, \dots, t_{10}$  mit zugehörigen Daten  $b_1, b_2, \dots, b_{10}$ .
- ▶ Modell einer gedämpften Schwingung

$$y(t; x_1, x_2, x_3, x_4) = x_1 e^{-x_2 t} \sin(x_3 t + x_4)$$

mit Parametern  $x_1, \dots, x_4$ .

# Beispiel 6.1

## Gegeben:

- ▶ 10 Messungen an den Punkten  $t_1, t_2, \dots, t_{10}$  mit zugehörigen Daten  $b_1, b_2, \dots, b_{10}$ .
- ▶ Modell einer gedämpften Schwingung

$$y(t; x_1, x_2, x_3, x_4) = x_1 e^{-x_2 t} \sin(x_3 t + x_4)$$

mit Parametern  $x_1, \dots, x_4$ .

## Gesucht:

- ▶ Parameter  $x_1, \dots, x_4$ , so dass die Summe der Fehlerquadrate

$$\sum_{i=1}^{10} (x_1 e^{-x_2 t_i} \sin(x_3 t_i + x_4) - b_i)^2 =: \|F(x)\|_2^2$$

minimal wird.

# Beispiel 6.1

## Gegeben:

- ▶ 10 Messungen an den Punkten  $t_1, t_2, \dots, t_{10}$  mit zugehörigen Daten  $b_1, b_2, \dots, b_{10}$ .
- ▶ Modell einer gedämpften Schwingung

$$y(t; x_1, x_2, x_3, x_4) = x_1 e^{-x_2 t} \sin(x_3 t + x_4)$$

mit Parametern  $x_1, \dots, x_4$ .

## Gesucht:

- ▶ Parameter  $x_1, \dots, x_4$ , so dass die Summe der Fehlerquadrate

$$\sum_{i=1}^{10} (x_1 e^{-x_2 t_i} \sin(x_3 t_i + x_4) - b_i)^2 =: \|F(x)\|_2^2$$

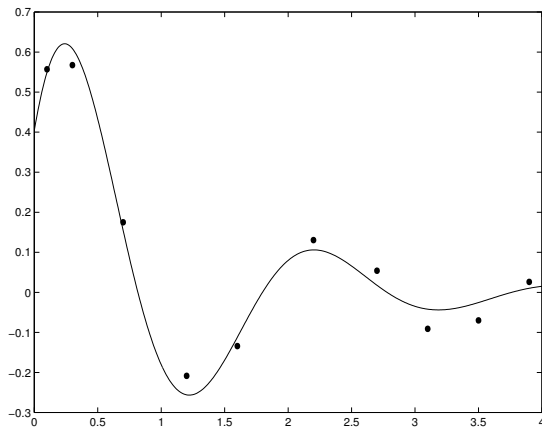
minimal wird.

Hierbei ist  $F: \mathbb{R}^4 \rightarrow \mathbb{R}^{10}$  nichtlinear in  $x_2, x_3, x_4$ :

$$F_i(x_1, x_2, x_3, x_4) = x_1 e^{-x_2 t_i} \sin(x_3 t_i + x_4) - b_i, \quad i = 1, \dots, 10.$$

# Beispiel 6.1.

Berechnete Lösung des nichtlinearen Ausgleichsproblems:



# Allgemeine Problemstellung

Definiert man allgemein die Abbildung ( $m > n$ )

$$F : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad F_i(x) := y(t_i; x) - b_i, \quad i = 1, \dots, m,$$

kann das **nichtlineare Ausgleichsproblem** wie folgt formuliert werden:

## Nichtlineares Ausgleichsproblem

Bestimme  $x^* \in U \subset \mathbb{R}^n$ , so dass

$$\|F(x^*)\|_2 = \min_{x \in U} \|F(x)\|_2,$$

oder, äquivalent,

$$\phi(x^*) = \min_{x \in U} \phi(x),$$

wobei

$$\phi : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \phi(x) := \frac{1}{2} \|F(x)\|_2^2 = \frac{1}{2} F(x)^T F(x).$$

# Bemerkungen

## Zur Erinnerung:

Die Funktion  $\phi(x) = \frac{1}{2}\|F(x)\|_2^2 = \frac{1}{2}F(x)^T F(x)$  hat in einem Punkt  $x^*$  ein **lokales Minimum**, wenn die folgenden zwei Bedingungen erfüllt sind:

# Bemerkungen

## Zur Erinnerung:

Die Funktion  $\phi(x) = \frac{1}{2}\|F(x)\|_2^2 = \frac{1}{2}F(x)^T F(x)$  hat in einem Punkt  $x^*$  ein **lokales Minimum**, wenn die folgenden zwei Bedingungen erfüllt sind:

1.  $\nabla\phi(x^*) = 0$  (d.h.  $x^*$  ist kritischer Punkt von  $\phi$ ),

# Bemerkungen

## Zur Erinnerung:

Die Funktion  $\phi(x) = \frac{1}{2}\|F(x)\|_2^2 = \frac{1}{2}F(x)^T F(x)$  hat in einem Punkt  $x^*$  ein **lokales Minimum**, wenn die folgenden zwei Bedingungen erfüllt sind:

1.  $\nabla\phi(x^*) = 0$  (d.h.  $x^*$  ist kritischer Punkt von  $\phi$ ),
2.  $\phi''(x^*) \in \mathbb{R}^{n \times n}$  ist symmetrisch positiv definit.

# Bemerkungen

## Zur Erinnerung:

Die Funktion  $\phi(x) = \frac{1}{2}\|F(x)\|_2^2 = \frac{1}{2}F(x)^T F(x)$  hat in einem Punkt  $x^*$  ein **lokales Minimum**, wenn die folgenden zwei Bedingungen erfüllt sind:

1.  $\nabla\phi(x^*) = 0$  (d.h.  $x^*$  ist kritischer Punkt von  $\phi$ ),
2.  $\phi''(x^*) \in \mathbb{R}^{n \times n}$  ist symmetrisch positiv definit.

Es läßt sich durch Nachrechnen bestätigen, dass

$$\nabla\phi(x) = F'(x)^T F(x),$$

$$\phi''(x) = F'(x)^T F'(x) + \sum_{i=1}^m F_i(x) F_i''(x),$$

mit **Jacobi-Matrix**  $F'(x) \in \mathbb{R}^{m \times n}$

und **Hesse-Matrix**  $F_i''(x) := \left( \frac{\partial^2 F_i(x)}{\partial x_j \partial x_k} \right)_{1 \leq j, k \leq n} \in \mathbb{R}^{n \times n}$

# Das Gauß-Newton-Verfahren

## Nichtlineares Ausgleichsproblem

Gegeben  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , bestimme  $x^* \in U \subset \mathbb{R}^n$ , so dass

$$\|F(x^*)\|_2 = \min_{x \in U} \|F(x)\|_2.$$

# Das Gauß-Newton-Verfahren

## Nichtlineares Ausgleichsproblem

Gegeben  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , bestimme  $x^* \in U \subset \mathbb{R}^n$ , so dass

$$\|F(x^*)\|_2 = \min_{x \in U} \|F(x)\|_2.$$

**Ansatz:**

# Das Gauß-Newton-Verfahren

## Nichtlineares Ausgleichsproblem

Gegeben  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , bestimme  $x^* \in U \subset \mathbb{R}^n$ , so dass

$$\|F(x^*)\|_2 = \min_{x \in U} \|F(x)\|_2.$$

### Ansatz:

1. Ersetze  $F(x)$  durch lineare Approximation (Taylor-Entwicklung)

# Das Gauß-Newton-Verfahren

## Nichtlineares Ausgleichsproblem

Gegeben  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , bestimme  $x^* \in U \subset \mathbb{R}^n$ , so dass

$$\|F(x^*)\|_2 = \min_{x \in U} \|F(x)\|_2.$$

### Ansatz:

1. Ersetze  $F(x)$  durch lineare Approximation (Taylor-Entwicklung)
2. Schrittweise Annäherung an  $x^*$  durch Lösung **linearer** Probleme in jedem Schritt

# Das Gauß-Newton-Verfahren

## Nichtlineares Ausgleichsproblem

Gegeben  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , bestimme  $x^* \in U \subset \mathbb{R}^n$ , so dass

$$\|F(x^*)\|_2 = \min_{x \in U} \|F(x)\|_2.$$

### Ansatz:

1. Ersetze  $F(x)$  durch lineare Approximation (Taylor-Entwicklung)
2. Schrittweise Annäherung an  $x^*$  durch Lösung **linearer** Probleme in jedem Schritt

### Zur Erinnerung:

# Das Gauß-Newton-Verfahren

## Nichtlineares Ausgleichsproblem

Gegeben  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , bestimme  $x^* \in U \subset \mathbb{R}^n$ , so dass

$$\|F(x^*)\|_2 = \min_{x \in U} \|F(x)\|_2.$$

### Ansatz:

1. Ersetze  $F(x)$  durch lineare Approximation (Taylor-Entwicklung)
2. Schrittweise Annäherung an  $x^*$  durch Lösung **linearer** Probleme in jedem Schritt

### Zur Erinnerung:

- Wir bezeichnen die Lösung am Iterationsschritt  $k$  mit

$$x^k = (x_1^k, \dots, x_n^k)^T \in \mathbb{R}^n.$$

# Das Gauß-Newton-Verfahren

## Nichtlineares Ausgleichsproblem

Gegeben  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , bestimme  $x^* \in U \subset \mathbb{R}^n$ , so dass

$$\|F(x^*)\|_2 = \min_{x \in U} \|F(x)\|_2.$$

### Ansatz:

1. Ersetze  $F(x)$  durch lineare Approximation (Taylor-Entwicklung)
2. Schrittweise Annäherung an  $x^*$  durch Lösung **linearer** Probleme in jedem Schritt

### Zur Erinnerung:

- ▶ Wir bezeichnen die Lösung am Iterationsschritt  $k$  mit

$$x^k = (x_1^k, \dots, x_n^k)^T \in \mathbb{R}^n.$$

- ▶ Taylor-Entwicklung

$$F(x) = F(x^k) + F'(x^k)(x - x^k) + \mathcal{O}(\|x - x^k\|_2^2).$$

# Das Gauß-Newton-Verfahren

## Ansatz:

Ersetze  $F(x)$  in  $\min_{x \in U} \|F(x)\|_2$ , durch lineare Approximation

$$\min_{x \in U} \|F(x^k) + F'(x^k) (x - x^k)\|_2.$$

# Das Gauß-Newton-Verfahren

## Ansatz:

Ersetze  $F(x)$  in  $\min_{x \in U} \|F(x)\|_2$ , durch lineare Approximation

$$\min_{x \in U} \underbrace{\|F(x^k) + F'(x^k)(x - x^k)\|_2}_{\in \mathbb{R}^m}.$$

Mit  $s := x - x^k$  (bzw.  $s^k = x^{k+1} - x^k$ ) ergibt sich das

### Lineare Ausgleichsproblem:

Finde  $s^k \in \mathbb{R}^n$  (mit minimaler 2-Norm), so dass

$$\|F'(x^k)s^k + F(x^k)\|_2 = \min_{s \in \mathbb{R}^n} \|F'(x^k)s + F(x^k)\|_2$$

# Das Gauß-Newton-Verfahren

## Ansatz:

Ersetze  $F(x)$  in  $\min_{x \in U} \|F(x)\|_2$ , durch lineare Approximation

$$\min_{x \in U} \left\| \underbrace{F(x^k)}_{\in \mathbb{R}^m} + \underbrace{F'(x^k)}_{\in \mathbb{R}^{m \times n}} (x - x^k) \right\|_2.$$

Mit  $s := x - x^k$  (bzw.  $s^k = x^{k+1} - x^k$ ) ergibt sich das

### Lineare Ausgleichsproblem:

Finde  $s^k \in \mathbb{R}^n$  (mit minimaler 2-Norm), so dass

$$\|F'(x^k)s^k + F(x^k)\|_2 = \min_{s \in \mathbb{R}^n} \|F'(x^k)s + F(x^k)\|_2$$

Anschließend berechnen wir

$$x^{k+1} = x^k + s^k.$$

# Das Gauß-Newton-Verfahren

## Ansatz:

Ersetze  $F(x)$  in  $\min_{x \in U} \|F(x)\|_2$ , durch lineare Approximation

$$\min_{x \in U} \underbrace{\|F(x^k) + F'(x^k) \underbrace{(x - x^k)}_{\in \mathbb{R}^n}\|_2}_{\substack{\in \mathbb{R}^m \\ \in \mathbb{R}^{m \times n}}}$$

Mit  $s := x - x^k$  (bzw.  $s^k = x^{k+1} - x^k$ ) ergibt sich das

### Lineare Ausgleichsproblem:

Finde  $s^k \in \mathbb{R}^n$  (mit minimaler 2-Norm), so dass

$$\|F'(x^k)s^k + F(x^k)\|_2 = \min_{s \in \mathbb{R}^n} \|F'(x^k)s + F(x^k)\|_2$$

Anschließend berechnen wir

$$x^{k+1} = x^k + s^k.$$

# Das Gauß-Newton-Verfahren

## Algorithmus 6.3 (Gauß-Newton)

Wähle Startwert  $x^0$ .

Für  $k = 0, 1, 2, \dots$ :

1. Berechne  $F(x^k)$ ,  $F'(x^k)$ .

2. Finde  $s^k \in \mathbb{R}^n$  mit minimaler 2-Norm, so dass

$$\|F'(x^k)s^k + F(x^k)\|_2 = \min_{s \in \mathbb{R}^n} \|F'(x^k)s + F(x^k)\|_2$$

3. Setze  $x^{k+1} = x^k + s^k$ .

# Das Gauß-Newton-Verfahren

## Algorithmus 6.3 (Gauß-Newton)

Wähle Startwert  $x^0$ .

Für  $k = 0, 1, 2, \dots$ :

1. Berechne  $F(x^k), F'(x^k)$ .
2. Finde  $s^k \in \mathbb{R}^n$  mit minimaler 2-Norm, so dass
$$\|F'(x^k)s^k + F(x^k)\|_2 = \min_{s \in \mathbb{R}^n} \|F'(x^k)s + F(x^k)\|_2$$
3. Setze  $x^{k+1} = x^k + s^k$ .

## Beachte

- ▶ Falls  $F'(x^k)$  vollen Rang hat, kann man den Zusatz “mit minimaler 2-Norm” weglassen.
- ▶ Schritt 2 erfordert die Lösung eines linearen Ausgleichsproblems.

# Bemerkungen

- ▶ Analogie zu nichtlinearen Gleichungssystemen: [Linearisierung](#)

# Bemerkungen

- ▶ Analogie zu nichtlinearen Gleichungssystemen: **Linearisierung**
- ▶ In einem **kritischen Punkt**  $x^*$  von  $\phi$  muss die **Ableitung**

$$\nabla \phi(x) = F'(x)^T F(x)$$

**gleich Null**  $\in \mathbb{R}^{m \times n}$  sein.

Als Abbruchkriterium für das Verfahren wird daher

$$\|F'(x^k)^T F(x^k)\|_2 \leq \varepsilon$$

häufig benutzt, wobei  $\varepsilon$  eine vorgegebene Toleranz ist.

# Bemerkungen

- ▶ Analogie zu nichtlinearen Gleichungssystemen: **Linearisierung**
- ▶ In einem **kritischen Punkt**  $x^*$  von  $\phi$  muss die **Ableitung**

$$\nabla \phi(x) = F'(x)^T F(x)$$

**gleich Null**  $\in \mathbb{R}^{m \times n}$  sein.

Als Abbruchkriterium für das Verfahren wird daher

$$\|F'(x^k)^T F(x^k)\|_2 \leq \varepsilon$$

häufig benutzt, wobei  $\varepsilon$  eine vorgegebene Toleranz ist.

- ▶ Der Erfolg des Gauß-Newton-Verfahrens hängt von der Wahl des Startwerts ab (vgl. Newton-Verfahren).

# Analyse der Gauß-Newton-Methode

Sei  $x^*$  ein kritischer Punkt von  $\phi$ , der in einer Umgebung  $U$  eindeutig ist.

Annahme:

$$\text{Rang}(F'(x)) = n \quad \text{für alle } x \in U .$$

Es gilt:

$$\nabla \phi(x) = 0 \Leftrightarrow x = \Phi(x), \text{ mit } \Phi(x) = x - (F'(x)^T F'(x))^{-1} \nabla \phi(x)$$

Fixpunktiteration:  $x_{k+1} = \Phi(x_k)$ .

# Analyse der Gauß-Newton-Methode

Sei  $x^*$  ein kritischer Punkt von  $\phi$ , der in einer Umgebung  $U$  eindeutig ist.

Annahme:

$$\text{Rang}(F'(x)) = n \quad \text{für alle } x \in U .$$

Es gilt:

$$\nabla \phi(x) = 0 \Leftrightarrow x = \Phi(x), \text{ mit } \Phi(x) = x - (F'(x)^T F'(x))^{-1} \nabla \phi(x)$$

Fixpunktiteration:  $x_{k+1} = \Phi(x_k)$ .

## Gauß-Newton als Fixpunktiteration

Für das Resultat  $x^{k+1} = x^k + s^k$  des Gauß-Newton-Verfahrens gilt

$$x^{k+1} = \Phi(x^k)$$

## Beispiel 6.4

Lösen Sie das nichtlineare Ausgleichsproblem

$$\min_{x \in \mathbb{R}} \|F(x)\|_2,$$

wobei

$$F(x) := \begin{pmatrix} a + r \cos x \\ r \sin x \end{pmatrix}, \text{ mit } a > r > 0, x \in [0, 2\pi).$$

## Beispiel 6.4

Lösen Sie das nichtlineare Ausgleichsproblem

$$\min_{x \in \mathbb{R}} \|F(x)\|_2,$$

wobei

$$F(x) := \begin{pmatrix} a + r \cos x \\ r \sin x \end{pmatrix}, \text{ mit } a > r > 0, x \in [0, 2\pi).$$

► Für die Jacobi-Matrix erhält man

$$F'(x) = r \begin{pmatrix} -\sin x \\ \cos x \end{pmatrix}, \quad F'(x)^T F'(x) = r^2.$$

## Beispiel 6.4

Lösen Sie das nichtlineare Ausgleichsproblem

$$\min_{x \in \mathbb{R}} \|F(x)\|_2,$$

wobei

$$F(x) := \begin{pmatrix} a + r \cos x \\ r \sin x \end{pmatrix}, \text{ mit } a > r > 0, x \in [0, 2\pi).$$

- Für die Jacobi-Matrix erhält man

$$F'(x) = r \begin{pmatrix} -\sin x \\ \cos x \end{pmatrix}, \quad F'(x)^T F'(x) = r^2.$$

- Außerdem ergibt sich

$$\phi(x) = \frac{1}{2} \|F(x)\|_2^2 = \frac{1}{2} (a^2 + 2ar \cos x + r^2)$$

und damit

$$\nabla \phi(x) = \phi'(x) = F'(x)^T F(x) = -ar \sin x.$$

## Beispiel 6.4

- Für die Iterationsfunktion zu  $F$  erhält man

$$\Phi(x) = x - (F'(x)^T F'(x))^{-1} \nabla \phi(x) = x + \frac{a}{r} \sin x$$

## Beispiel 6.4

- Für die Iterationsfunktion zu  $F$  erhält man

$$\Phi(x) = x - (F'(x)^T F'(x))^{-1} \nabla \phi(x) = x + \frac{a}{r} \sin x$$

- Es gibt zwei kritische Punkte von  $\phi$

# Beispiel 6.4

- Für die Iterationsfunktion zu  $F$  erhält man

$$\Phi(x) = x - (F'(x)^T F'(x))^{-1} \nabla \phi(x) = x + \frac{a}{r} \sin x$$

- Es gibt zwei kritische Punkte von  $\phi$

$$x^* = 0 \quad (\text{lokales Maximum}),$$

$$x^* = \pi \quad (\text{lokales Minimum}).$$

## Beispiel 6.4

- Für die Iterationsfunktion zu  $F$  erhält man

$$\Phi(x) = x - (F'(x)^T F'(x))^{-1} \nabla \phi(x) = x + \frac{a}{r} \sin x$$

- Es gibt zwei kritische Punkte von  $\phi$

$$x^* = 0 \quad (\text{lokales Maximum}),$$

$$x^* = \pi \quad (\text{lokales Minimum}).$$

- In den kritischen Punkten  $x^* = 0$ ,  $x^* = \pi$  gilt

$$|\Phi'(x^*)| = |1 + \frac{a}{r} \cos x^*|.$$

## Beispiel 6.4

- Für die Iterationsfunktion zu  $F$  erhält man

$$\Phi(x) = x - (F'(x)^T F'(x))^{-1} \nabla \phi(x) = x + \frac{a}{r} \sin x$$

- Es gibt zwei kritische Punkte von  $\phi$

$$x^* = 0 \quad (\text{lokales Maximum}),$$

$$x^* = \pi \quad (\text{lokales Minimum}).$$

- In den kritischen Punkten  $x^* = 0$ ,  $x^* = \pi$  gilt

$$|\Phi'(x^*)| = |1 + \frac{a}{r} \cos x^*|.$$

und damit

$$|\Phi'(0)| = \frac{a+r}{r} > 1 \quad (\text{im lokalen Max.})$$

$$|\Phi'(\pi)| = \frac{a-r}{r} = \frac{a}{r} - 1 \quad (\text{im lokalen Min.})$$

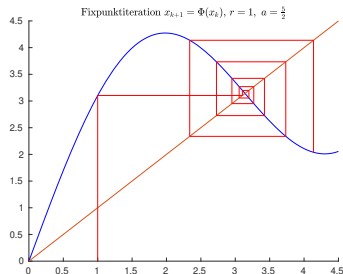
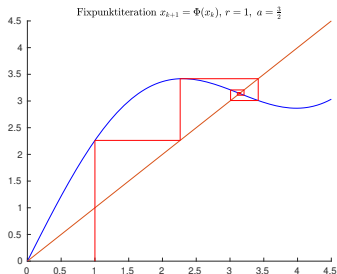
# Beispiel 6.4

Resultate der Fixpunktiteration  $x_{k+1} = \Phi(x_k)$ . Parameterwert  $r = 1$ .

$$a = \frac{3}{2}, \phi'(0) = \frac{5}{2}, \phi'(\pi) = \frac{1}{2}$$

$$\phi'(\pi) = \frac{3}{2}$$

$$a = \frac{5}{2}, \phi'(0) = \frac{7}{2},$$



## Beispiel 6.4.

Das Gauß-Newton-Verfahren hat in diesem Beispiel folgende Eigenschaften:

1. Das lokale Maximum ist abstoßend

## Beispiel 6.4.

Das Gauß-Newton-Verfahren hat in diesem Beispiel folgende Eigenschaften:

1. Das **lokale Maximum** ist abstoßend
2. Die Methode ist **linear** konvergent in einer Umgebung des **lokalen Minimums** (wenn  $a < 2r$ ), oder

## Beispiel 6.4.

Das Gauß-Newton-Verfahren hat in diesem Beispiel folgende Eigenschaften:

1. Das **lokale Maximum** ist **abstoßend**
2. Die Methode ist **linear** konvergent in einer Umgebung des **lokalen Minimums** (wenn  $a < 2r$ ), oder
3. das **lokale Minimum** ist **auch abstoßend** (wenn  $a > 2r$ ).

Man kann zeigen, dass ähnliche Eigenschaften in einem allgemeinen Rahmen gültig sind.

# Allgemeine Analyse

Sei  $A \in \mathbb{R}^{n \times n}$  die symmetrisch positiv definite Matrix so dass

$$A^2 = F'(x^*)^T F'(x^*).$$

Wir definieren

$$K := -A^{-1} \left( \sum_{i=1}^m \frac{F_i(x^*)}{\|F(x^*)\|_2} F_i''(x^*) \right) A^{-1}.$$

Weil  $K$  eine symmetrische Matrix ist, sind alle Eigenwerte von  $K$  reell.

## Lemma 6.6

Wenn  $x^*$  ein lokales Maximum oder ein Sattelpunkt von  $\phi$  ist, muss

$$\rho(K) \|F(x^*)\|_2 \geq 1$$

gelten, wobei  $\rho(K)$  den Spektralradius von  $K$  bezeichnet.

# Allgemeine Analyse

## Folgerung 6.7

Für die allgemeine Gauß-Newton-Iterationsfunktion  $\Phi$  gilt

$$\|\Phi'(x^*)\|_A = \rho(K) \|F(x^*)\|_2 ,$$

$$\|\Phi'(x^*)\| \geq \rho(K) \|F(x^*)\|_2, \quad \text{für jede Operatornorm } \|\cdot\|$$

# Allgemeine Analyse

## Folgerung 6.7

Für die allgemeine Gauß-Newton-Iterationsfunktion  $\Phi$  gilt

$$\|\Phi'(x^*)\|_A = \rho(K) \|F(x^*)\|_2 ,$$

$$\|\Phi'(x^*)\| \geq \rho(K) \|F(x^*)\|_2, \quad \text{für jede Operatornorm } \|\cdot\|$$

Im Normalfall ist  $F(x^*) \neq 0$ ,  $K \neq 0$  und deshalb  $\Phi'(x^*) \neq 0$ .

Falls die Gauß-Newton-Methode konvergiert, ist die **Konvergenz im allgemeinen nicht schneller als linear**.

# Allgemeine Analyse

Wenn der kritische Punkt  $x^*$  ein **lokales Maximum** oder ein **Sattelpunkt** von  $\phi$  ist, gilt (Lemma 6.6)  $\rho(K)\|F(x^*)\|_2 \geq 1$ .

Kritische Punkte von  $F$ , die **kein** lokales Minimum sind, sind für das Gauß-Newton-Verfahren **abstoßend**.

Beachte: dies ist eine günstige Eigenschaft, da ein (lokales) Minimum gesucht wird und das Verfahren somit “**falsche**” kritische Punkt ignoriert.

# Allgemeine Analyse

Die Größe  $\rho(K)\|F(x^*)\|_2$  ist entscheidend für die lokale Konvergenz des Gauß-Newton-Verfahrens.

Für ein lokales Minimum  $x^*$  der Funktion  $\phi$  ist die **lokale Konvergenz des Gauß-Newton-Verfahrens gesichert**, falls die Bedingung

$$\rho(K)\|F(x^*)\|_2 < 1$$

erfüllt ist.

# Allgemeine Analyse

Die Größe  $\rho(K)\|F(x^*)\|_2$  ist entscheidend für die lokale Konvergenz des Gauß-Newton-Verfahrens.

Für ein lokales Minimum  $x^*$  der Funktion  $\phi$  ist die **lokale Konvergenz des Gauß-Newton-Verfahrens gesichert**, falls die Bedingung

$$\rho(K)\|F(x^*)\|_2 < 1$$

erfüllt ist.

Sei  $x^*$  ein lokales Minimum von  $\phi$ , wofür gilt:

$$\rho(K)\|F(x^*)\|_2 > 1.$$

Dann ist  $\|\Phi'(x^*)\| > 1$  für jede Operatornorm  $\|\cdot\|$ .

Ein lokales Minimum von  $\phi$  **kann** für die Gauß-Newton-Methode also abstoßend sein.

## Beispiel 6.8

Das Gauß-Newton-Verfahren angewandt auf das Problem der gedämpften Schwingung in Beispiel 6.1.

$k$	$\ F(x^k)\ _2$	$\ \nabla\phi(x^k)\ _2$	$\ \nabla\phi(x^k)\ _2/\ \nabla\phi(x^{k-1})\ _2$
0	0.35035332090089	1.45e-01	-
1	0.34106434131008	1.33e-01	0.91
2	0.22208131421995	4.88e-02	0.37
3	0.16802866234936	1.02e-01	2.08
4	0.09190056278958	1.80e-01	0.18
5	0.08902339976144	1.18e-03	0.07
6	0.08895515308450	3.81e-04	0.32
7	0.08894991006370	1.15e-04	0.30
8	0.08894937563528	4.07e-05	0.35
9	0.08894931422207	1.38e-05	0.34
10	0.08894930687791	4.85e-06	0.35
11	0.08894930599062	1.68e-06	0.35
12	0.08894930588306	5.87e-07	0.35

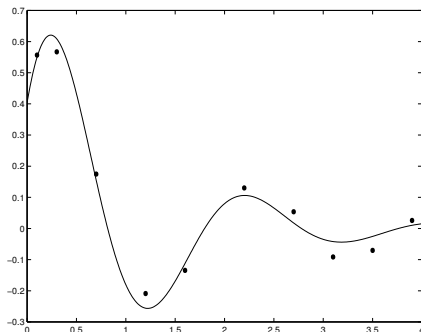
In der letzten Spalte der Tabelle sieht man das lineare Konvergenzverhalten des Gauß-Newton-Verfahrens.

## Beispiel 6.8

Die berechneten Parameterwerte  $x^{12} \approx x^*$  aus dem 12 Iterationsschritt liefern die Lösung

$$y(t; x^*) = x_1^* e^{-x_2^* t} \sin(x_3^* t + x_4^*),$$

die im folgenden Plot dargestellt ist.



# Zusammenfassung

- ▶ Beim Gauß-Newton-Verfahren wird das **nichtlineare Ausgleichsproblem** über eine Folge **linearer Ausgleichsprobleme** gelöst.

# Zusammenfassung

- ▶ Beim Gauß-Newton-Verfahren wird das **nichtlineare Ausgleichsproblem** über eine Folge **linearer Ausgleichsprobleme** gelöst.
- ▶ Lokale Konvergenz im allgemeinen nur **1. Ordnung**, falls  $F(x^*) = 0$ , sogar von **2. Ordnung**.

# Zusammenfassung

- ▶ Beim Gauß-Newton-Verfahren wird das **nichtlineare Ausgleichsproblem** über eine Folge **linearer Ausgleichsprobleme** gelöst.
- ▶ Lokale Konvergenz im allgemeinen nur **1. Ordnung**, falls  $F(x^*) = 0$ , sogar von **2. Ordnung**.
- ▶ Es kann lokale **Divergenz** auftreten.

# Zusammenfassung

- ▶ Beim Gauß-Newton-Verfahren wird das **nichtlineare Ausgleichsproblem** über eine Folge **linearer Ausgleichsprobleme** gelöst.
- ▶ Lokale Konvergenz im allgemeinen nur **1. Ordnung**, falls  $F'(x^*) = 0$ , sogar von **2. Ordnung**.
- ▶ Es kann lokale **Divergenz** auftreten.
- ▶ Matrix  $F'(x^k)$  kann  $\text{Rang} < n$  haben.