

Numerische Methoden

Kapitel 1: Einleitung

Heinrich Voss

voss@tu-harburg.de

Hamburg University of Technology
Institute for Numerical Simulation



Aufgabe der Numerischen Mathematik ist, Algorithmen (d.h. Rechenvorschriften) für die näherungsweise numerische Lösung mathematischer Probleme der Naturwissenschaften, Technik, Ökonomie u.s.w. bereitzustellen und zu diskutieren.

Aufgabe der Numerischen Mathematik ist, Algorithmen (d.h. Rechenvorschriften) für die näherungsweise numerische Lösung mathematischer Probleme der Naturwissenschaften, Technik, Ökonomie u.s.w. bereitzustellen und zu diskutieren.

Gesichtspunkte bei der Bewertung eines Algorithmus (und beim Vergleich von Algorithmen) sind der Aufwand (unter Berücksichtigung der Hardware, früher einfach: Zahl der Operationen), Speicherplatzbedarf und eine Fehleranalyse.

Fehlertypen

Man unterscheidet drei Typen von Fehlern nach ihren Quellen:

Fehlertypen

Man unterscheidet drei Typen von Fehlern nach ihren Quellen:

1. **Datenfehler:** Die Eingangsdaten einer Aufgabe können fehlerhaft sein, wenn sie etwa aus vorhergehenden Rechnungen, physikalischen Messungen oder empirischen Untersuchungen stammen.

Fehlertypen

Man unterscheidet drei Typen von Fehlern nach ihren Quellen:

1. **Datenfehler**: Die Eingangsdaten einer Aufgabe können fehlerhaft sein, wenn sie etwa aus vorhergehenden Rechnungen, physikalischen Messungen oder empirischen Untersuchungen stammen.
2. **Verfahrensfehler**: Dies sind Fehler, die dadurch entstehen, dass man ein Problem diskretisiert. (z.B. eine Differentialgleichung durch eine Differenzengleichung ersetzt) oder ein Iterationsverfahren nach endlich vielen Schritten abbricht.

Fehlertypen

Man unterscheidet drei Typen von Fehlern nach ihren Quellen:

1. **Datenfehler:** Die Eingangsdaten einer Aufgabe können fehlerhaft sein, wenn sie etwa aus vorhergehenden Rechnungen, physikalischen Messungen oder empirischen Untersuchungen stammen.
2. **Verfahrensfehler:** Dies sind Fehler, die dadurch entstehen, dass man ein Problem diskretisiert. (z.B. eine Differentialgleichung durch eine Differenzengleichung ersetzt) oder ein Iterationsverfahren nach endlich vielen Schritten abbricht.
3. **Rundungsfehler:** Bei der Ausführung der Rechenoperationen auf einer Rechenanlage entstehen Fehler, da das Ergebnis (aber auch schon alle Zwischenergebnisse) nur im Rahmen eines begrenzten Zahlbereiches (sog. Maschinenzahlen) dargestellt werden kann, also gerundet werden muss.

Konditionsproblem

Die Frage, wie sich Datenfehler auf die Lösung einer Aufgabe auswirken, nennt man das Konditionsproblem der Aufgabe.

Konditionsproblem

Die Frage, wie sich Datenfehler auf die Lösung einer Aufgabe auswirken, nennt man das Konditionsproblem der Aufgabe.

Bewirken kleine Eingangsfehler auch nur kleine Ergebnisfehler, so nennt man das Problem **gut konditioniert**, anderenfalls **schlecht konditioniert**.

Konditionsproblem

Die Frage, wie sich Datenfehler auf die Lösung einer Aufgabe auswirken, nennt man das Konditionsproblem der Aufgabe.

Bewirken kleine Eingangsfehler auch nur kleine Ergebnisfehler, so nennt man das Problem **gut konditioniert**, anderenfalls **schlecht konditioniert**.

Die Kondition eines Problems hängt nicht nur von der Aufgabenstellen (z.B. “Lösung eines linearen Gleichungssystems”), sondern auch von den Eingangsdaten ab (z.B. Koeffizienten der Matrix)

Konditionsproblem cnt.

Wir beschreiben das zu lösende Problem durch eine Funktion

$$f : X \rightarrow Y, \quad X, Y \text{ normierte Vektorräume,}$$

die an einer Stelle $x \in X$ ausgewertet werden muss.

Konditionsproblem cnt.

Wir beschreiben das zu lösende Problem durch eine Funktion

$$f : X \rightarrow Y, \quad X, Y \text{ normierte Vektorräume,}$$

die an einer Stelle $x \in X$ ausgewertet werden muss.

Sei δx eine kleine Störung von x . Dann ist $\delta f := f(x + \delta x) - f(x)$ der absolute Fehler. Die **absolute Konditionszahl** $\hat{\kappa} = \hat{\kappa}(x)$ ist definiert durch

$$\hat{\kappa} := \lim_{\delta \rightarrow 0} \sup_{\|\delta x\| \leq \delta} \frac{\|\delta f\|}{\|\delta x\|}.$$

Konditionsproblem cnt.

Wir beschreiben das zu lösende Problem durch eine Funktion

$$f : X \rightarrow Y, \quad X, Y \text{ normierte Vektorräume,}$$

die an einer Stelle $x \in X$ ausgewertet werden muss.

Sei δx eine kleine Störung von x . Dann ist $\delta f := f(x + \delta x) - f(x)$ der absolute Fehler. Die **absolute Konditionszahl** $\hat{\kappa} = \hat{\kappa}(x)$ ist definiert durch

$$\hat{\kappa} := \lim_{\delta \rightarrow 0} \sup_{\|\delta x\| \leq \delta} \frac{\|\delta f\|}{\|\delta x\|}.$$

Ist f differenzierbar in x , so gilt $f(x + \delta x) = f(x) + J(x)\delta x + o(\|\delta x\|)$ mit der Jacobi Matrix $J(x)$, und es folgt

$$\hat{\kappa} = \|J(x)\|.$$

Beispiel

Das lineare Gleichungssystem

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad a \in \mathbb{R}$$

besitzt die Lösung

$$x_0 = 1 \quad , \quad y_0 = 0 \quad .$$

Beispiel

Das lineare Gleichungssystem

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad a \in \mathbb{R}$$

besitzt die Lösung

$$x_0 = 1 \quad , \quad y_0 = 0 \quad .$$

Das gestörte Gleichungssystem

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ \delta \end{pmatrix}$$

hat die Lösung

$$x_\delta = 1 - \delta a \quad , \quad y_\delta = \delta \quad .$$

Beispiel cnt.

Damit gilt

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} x_\delta \\ y_\delta \end{pmatrix} = \delta \begin{pmatrix} -a \\ 1 \end{pmatrix}.$$

Beispiel cnt.

Damit gilt

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} x_\delta \\ y_\delta \end{pmatrix} = \delta \begin{pmatrix} -a \\ 1 \end{pmatrix}.$$

Änderungen der rechten Seite $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ in der zweiten Komponente werden also mit dem Faktor $\sqrt{1+a^2}$ (bzgl. der Euklidischen Norm) verstärkt. Damit ist das Problem für kleine $|a|$ die absolute Konditionszahl klein und für große $|a|$ groß.

□

Konditionszahl

Wichtiger als die absolute Konditionszahl ist die relative Konditionszahl κ (insbesondere wenn man mit Gleitpunktzahlen arbeitet)

$$\kappa := \lim_{\delta \rightarrow 0} \sup_{\|\delta x\| \leq \delta} \left(\frac{\|\delta f\|}{\|f(x)\|} / \frac{\|\delta x\|}{\|x\|} \right),$$

Konditionszahl

Wichtiger als die absolute Konditionszahl ist die relative Konditionszahl κ (insbesondere wenn man mit Gleitpunktzahlen arbeitet)

$$\kappa := \lim_{\delta \rightarrow 0} \sup_{\|\delta x\| \leq \delta} \left(\frac{\|\delta f\|}{\|f(x)\|} / \frac{\|\delta x\|}{\|x\|} \right),$$

und im differenzierbaren Fall

$$\kappa = \frac{\|J(x)\|}{\|f(x)\| / \|x\|}.$$

Konditionszahl

Wichtiger als die absolute Konditionszahl ist die relative Konditionszahl κ (insbesondere wenn man mit Gleitpunktzahlen arbeitet)

$$\kappa := \lim_{\delta \rightarrow 0} \sup_{\|\delta x\| \leq \delta} \left(\frac{\|\delta f\|}{\|f(x)\|} / \frac{\|\delta x\|}{\|x\|} \right),$$

und im differenzierbaren Fall

$$\kappa = \frac{\|J(x)\|}{\|f(x)\| / \|x\|}.$$

Im Beispiel gilt

$$\lim_{\delta \rightarrow 0} \sup_{\|\delta x\| \leq \delta} \left(\frac{\|\delta f\|}{\|f(x)\|} / \frac{\|\delta x\|}{\|x\|} \right) = \lim_{\delta \rightarrow 0} \sup_{\|\delta x\| \leq \delta} \left(\frac{|\delta(1+a)|}{1} / \frac{|\delta|}{1} \right) = 1 + |a|,$$

Beispiel

Berechne $\sqrt{x}$ für ein $x > 0$.

Beispiel

Berechne $\sqrt{x}$ für ein $x > 0$.

Die Jacobi Matrix von $f : x \mapsto \sqrt{x}$ ist die Ableitung $J = f' = \frac{1}{2\sqrt{x}}$.

Beispiel

Berechne $\sqrt{x}$ für ein $x > 0$.

Die Jacobi Matrix von $f : x \mapsto \sqrt{x}$ ist die Ableitung $J = f' = \frac{1}{2\sqrt{x}}$.

Daher folgt

$$\kappa = \frac{\|J(x)\|}{\|f(x)\|/\|x\|} = \frac{1/2\sqrt{x}}{\sqrt{x}/x} = \frac{1}{2},$$

und das Problem ist gut konditioniert.

Beispiel

Berechne $f(x) = x_1 - x_2$ für einen Vektor $x \in \mathbb{C}^2$.

Beispiel

Berechne $f(x) = x_1 - x_2$ für einen Vektor $x \in \mathbb{C}^2$.

Die Jacobi Matrix von $f : (x_1, x_2) \mapsto x_1 - x_2$ ist

$$J(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{pmatrix} = (1 \quad -1).$$

Beispiel

Berechne $f(x) = x_1 - x_2$ für einen Vektor $x \in \mathbb{C}^2$.

Die Jacobi Matrix von $f : (x_1, x_2) \mapsto x_1 - x_2$ ist

$$J(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{pmatrix} = (1 \quad -1).$$

Daher folgt unter Benutzung der ∞ -norm(, wenn Sie mir $\|J\| = 2$ glauben,)

$$\kappa = \frac{\|J(x)\|}{\|f(x)\|/\|x\|} = \frac{2}{|x_1 - x_2|/\max\{|x_1|, |x_2|\}},$$

und das Problem ist schlecht konditioniert, wenn $x_1 \approx x_2$ gilt, sonst ist es gut konditioniert.

Konditionsproblem

Ein numerisches Verfahren heißt **gut konditioniert** (**numerisch stabil**), wenn die gelieferte Lösung eines gegebenen Problems die exakte Lösung eines Problems ist, das aus dem ursprünglichen Problem durch geringe Änderung der Eingangsdaten hervorgeht.

Konditionsproblem

Ein numerisches Verfahren heißt **gut konditioniert** (**numerisch stabil**), wenn die gelieferte Lösung eines gegebenen Problems die exakte Lösung eines Problems ist, das aus dem ursprünglichen Problem durch geringe Änderung der Eingangsdaten hervorgeht.

Anderenfalls heißt das numerische Verfahren **schlecht konditioniert** (oder **numerisch instabil**)

Beispiel

Bestimme $\int_0^1 \frac{x^{20}}{10+x} dx$

Beispiel

Bestimme $\int_0^1 \frac{x^{20}}{10+x} dx$

Für

$$y_n := \int_0^1 \frac{x^n}{10+x} dx$$

gilt

$$y_n + 10y_{n-1} = \int_0^1 \frac{x^n + 10x^{n-1}}{x+10} dx = \int_0^1 x^{n-1} dx = \frac{1}{n} \quad (*),$$

$$y_0 = \int_0^1 \frac{dx}{10+x} = \ln(1.1) .$$

Beispiel

Wertet man die Differenzenformel $y_n = \frac{1}{n} - 10y_{n-1}$ für $n = 1, \dots, 20$ aus, so erhält man die Werte der folgenden Tabelle.

Beispiel

Wertet man die Differenzenformel $y_n = \frac{1}{n} - 10y_{n-1}$ für $n = 1, \dots, 20$ aus, so erhält man die Werte der folgenden Tabelle.

Obwohl das Problem, das Integral zu berechnen, gut konditioniert ist, erhält man ein unbrauchbares Resultat. Das Verfahren ist also instabil.

Beispiel

Wertet man die Differenzenformel $y_n = \frac{1}{n} - 10y_{n-1}$ für $n = 1, \dots, 20$ aus, so erhält man die Werte der folgenden Tabelle.

Obwohl das Problem, das Integral zu berechnen, gut konditioniert ist, erhält man ein unbrauchbares Resultat. Das Verfahren ist also instabil.

Löst man (*) nach y_{n-1} auf:

$$y_{n-1} = 0.1 \left(\frac{1}{n} - y_n \right)$$

und startet man mit der groben Näherung $y_{30} = 0$, so erhält man $y_{20}, \dots, y_0$ mit einer Genauigkeit von wenigstens 10 gültigen Stellen. $\square$

Beispiel

n	vorwärts	rückwärts
0	9.53101798043249E-0002	9.53101798043249E-0002
1	4.68982019567514E-0002	4.68982019567514E-0002
2	3.10179804324860E-0002	3.10179804324860E-0002
3	2.31535290084733E-0002	2.31535290084733E-0002
4	1.84647099152673E-0002	1.84647099152671E-0002
5	1.53529008473272E-0002	1.53529008473289E-0002
6	1.31376581933944E-0002	1.31376581933773E-0002
7	1.14805609231986E-0002	1.14805609233700E-0002
8	1.01943907680135E-0002	1.01943907663000E-0002
9	9.16720343097581E-0003	9.16720344811137E-0003
10	8.32796569024192E-0003	8.32796551888631E-0003
11	7.62943400667172E-0003	7.62943572022779E-0003
12	7.03899326661614E-0003	7.03897613105546E-0003
13	6.53314425691553E-0003	6.53331561252233E-0003
14	6.09712885941609E-0003	6.09541530334817E-0003
15	5.69537807250574E-0003	5.71251363318499E-0003
16	5.54621927494255E-0003	5.37486366815006E-0003
17	3.36133666233920E-0003	5.07489273026414E-0003
18	2.19421889321635E-0002	4.80662825291418E-0003
19	-1.66790310374267E-0001	4.56529641822660E-0003
20	1.71790310374267E+0000	4.34703581773402E-0003
21		4.14868944170746E-0003
22		3.96765103747089E-0003
23		3.80175049485636E-0003
24		3.64916171810310E-0003
25		3.50838281896903E-0003

Stabilität

Ein Algorithmus zur Lösung des Problems $f : X \rightarrow Y$ ist eine weitere Abbildung $\tilde{f} : X \rightarrow Y$.

Stabilität

Ein Algorithmus zur Lösung des Problems $f : X \rightarrow Y$ ist eine weitere Abbildung $\tilde{f} : X \rightarrow Y$.

Führt man einen Algorithmus in Gleitpunktarithmetik aus, so ist das Beste, was man erreichen kann, dass der relative Fehler von der Ordnung der Rechengenauigkeit ist:

$$\frac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|} = O(\epsilon_{\text{machine}}).$$

In diesem Fall nennt man den Algorithmus exakt.

Stabilität

Ein Algorithmus zur Lösung des Problems $f : X \rightarrow Y$ ist eine weitere Abbildung $\tilde{f} : X \rightarrow Y$.

Führt man einen Algorithmus in Gleitpunktarithmetik aus, so ist das Beste, was man erreichen kann, dass der relative Fehler von der Ordnung der Rechengenauigkeit ist:

$$\frac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|} = O(\epsilon_{\text{machine}}).$$

In diesem Fall nennt man den Algorithmus exakt.

Ist das Problem f schlecht konditioniert, so kann man nicht erwarten, dass ein Algorithmus dieses exakt löst. Auftretende Rundungsfehler werden ja (möglicherweise) um die Kondition des Problems verstärkt und machen das Ergebnis unbrauchbar.

Stabilität

Ein numerisches Verfahren heißt (**rückwärts stabil**), wenn die gelieferte Lösung eines gegebenen Problems die exakte Lösung eines Problems ist, das aus dem ursprünglichen Problem durch geringe Änderung der Eingangsdaten hervorgeht,

Stabilität

Ein numerisches Verfahren heißt (**rückwärts stabil**), wenn die gelieferte Lösung eines gegebenen Problems die exakte Lösung eines Problems ist, das aus dem ursprünglichen Problem durch geringe Änderung der Eingangsdaten hervorgeht,

d.h. für alle $x \in X$ gilt

$$\tilde{f}(x) = f(\tilde{x}) \quad \text{für ein } \tilde{x} \text{ mit } \frac{\|x - \tilde{x}\|}{\|x\|} = O(\epsilon_{\text{machine}}).$$

Stabilität

Ein numerisches Verfahren heißt (**rückwärts stabil**), wenn die gelieferte Lösung eines gegebenen Problems die exakte Lösung eines Problems ist, das aus dem ursprünglichen Problem durch geringe Änderung der Eingangsdaten hervorgeht,

d.h. für alle $x \in X$ gilt

$$\tilde{f}(x) = f(\tilde{x}) \quad \text{für ein } \tilde{x} \text{ mit } \frac{\|x - \tilde{x}\|}{\|x\|} = O(\epsilon_{\text{machine}}).$$

Ein rückwärts stabiler Algorithmus gibt also die exakte Antwort auf die nahezu richtige Frage.