

1 LINEARE GLEICHUNGSSYSTEME

Im folgenden werden wir uns mit der Gaußschenfootnote¹ Eliminationsmethode zur Lösung von linearen Gleichungssystemen beschäftigen, bei dem wir voraussetzen, dass es eine eindeutige Lösung gibt und die Anzahl an Unbekannten und Gleichungen gleich ist. Obwohl mit dem Namen Gauss verbunden, wurde schon von Lagrange² 1759 eine Methode veröffentlicht, welches die grundlegenden Ideen enthält. Gauss veröffentlichte das von ihm im Rahmen seiner Entwicklung und Anwendung der Methode der kleinsten Quadrate von modifizierte Verfahren mitunter in in *Theoria motus corporum coelestium in sectionibus conicis solem ambientium* (Theorie der Bewegung der Himmelskörper, die in Kegelschnitten die Sonne umlaufen), das 1809 erschien. In realen Anwendungen, war das Verfahren von Cramer wegen des großen Rechenaufwands nicht anwendbar. Gauss entwickelte daher ein systematisches eliminieren für Gleichungssysteme, eine Methode bei der gewisse Gleichungen mit geeigneten Faktoren multipliziert und dann addiert werden. Die nun unter dem Namen Gauss-Verfahren bekannte Methode, ist mit dem vor über 2100 Jahren in China verwendeten Verfahren identisch. In der frühen Han-Zeit entstand in das chinesische Rechenbuch „Neun Bücher arithmetischer Technik“ Dabei handelt es sich um eine Zusammenstellung der in den vorhergehenden Jahrhunderten entwickelten Erkenntnisse in Gestalt einer Sammlung praktischer Aufgaben mit ihren jeweiligen Lösungsgängen. Hieraus ist das folgende Problem entnommen.

Drei Bauern handeln auf einem Markt Rinder, Schafe und Schweine. Der erste Bauer verkauft 2 Rinder und 5 Schafe und kauft 13 Schweine, wobei 1000 Geldstücke übrigbleiben. Der zweite verkauft 3 Rinder und 3 Schweine, wofür er genau 9 Schafe bekommt. Der dritte Bauer verkauft 6 Schafe und 8 Schweine und kauft 5 Rinder, wobei er noch 600 Geldstücke drauflegen muss. Wie hoch ist der Preis eines Rinds, eines Schafs und eines Schweins?

aus „Neun Bücher arithmetischer Technik“
(1. Jahrhunderts n. Chr)

1.1 KURZE MOTIVATION VON LINEAREN GLEICHUNGSSYSTEMEN

Betrachten wir doch noch zwei weitere motivierende Beispiele für lineare Gleichungssysteme.

1.1.1 Netzwerkanalyse

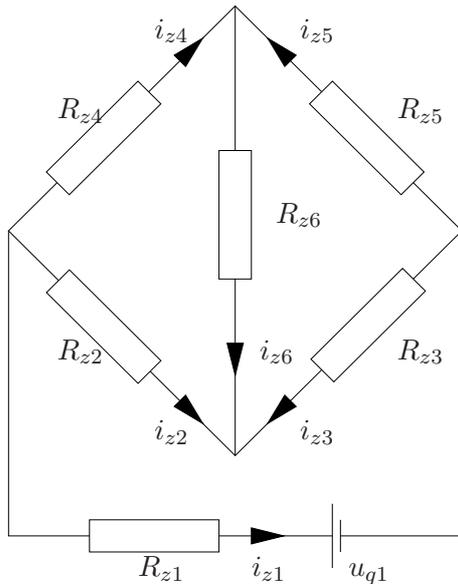
Die Maschen- und die Knotenanalyse sind die formalisierten Varianten der Anwendung der Kirchhoffschen Gesetze auf umfangreichere elektrische Netzwerk. In beiden Fällen wird die Schaltung in Untereinheiten zerlegt, allerdings nicht auf Grund eines Sinnzusammenhanges sondern auf der Basis der topologischen Merkmale Masche und Knoten.

Das Aufstellen des vollständigen Gleichungssystems ist etwas mühsam und in der Regel auch nicht erforderlich. Die Maschenstromanalyse ist eins von zwei gebräuchlichen abkürzenden Verfahren, welches wir hier genauer betrachten wollen.

Die Maschenanalyse reduziert die Zahl der zu lösenden Gleichungen um die Knotengleichungen. Anstelle der Zweigströme werden dazu Maschen- oder Kreisströme eingeführt, durch deren Überlagerung sich die Zweigströme ergeben.

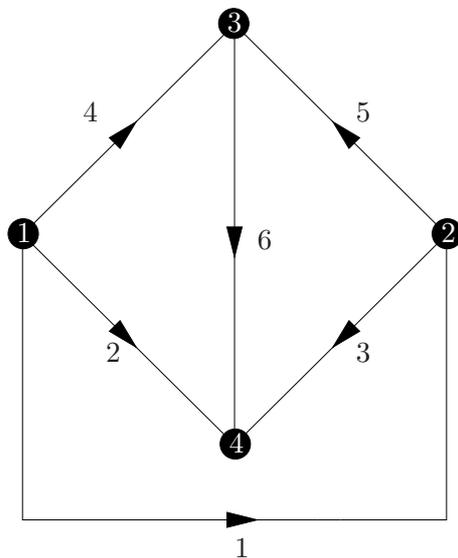
¹Carl Friedrich Gauß, 1777 - 1855

²Lagrange, 1777 - 1855



Es seien $R_{z\mu}$ die ohmschen Widerstände auf den Zweigen und $i_{z\mu}$ die Zweigströme des Zweigs μ . Der Spannungsverlust ohne zusätzliche Spannungsquelle ist nach dem ohmschen Gesetz somit $u_{z\mu} = R_{z\mu} i_{z\mu}$.

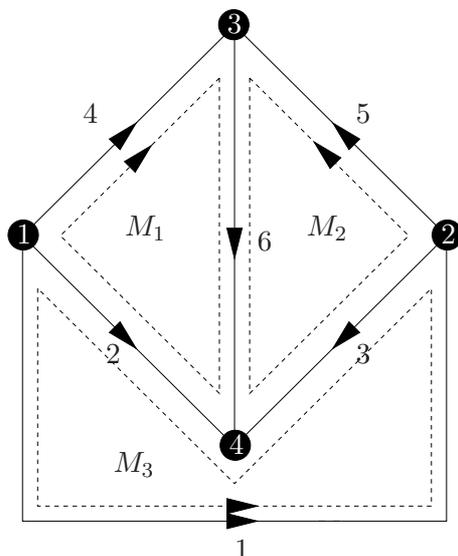
Ziel ist es nun, aus der gegebenen Quellenspannung u_{q1} und den Zweigwiderständen $R_{z\mu}$ die Zweigströme $i_{z\mu}$ und -spannungen $u_{z\mu}$ zu berechnen. Wieso führt dies auf ein lineares Gleichungssystem?



Legen wir nun das Bezugssystem fest.

- Nummerierung der Knoten $1, \dots, k$ (hier $k = 4$)
- Nummerierung der Zweige $1, \dots, z$ (hier $z = 6$)
- Festlegen der Stromrichtung

Desweiteren ermitteln wir linear unabhängige Maschen, d.h. geschlossene Wege im Graphen.



Das 2. Kirchhoffsche Gesetz oder der als Maschenregel bekannte Satz besagt, dass alle Teilspannungen eines Umlaufs bzw. einer Masche in einem elektrischen Netzwerk sich zu null addieren.

↓ Maschen/Zweige →

	1	2	3	4	5	6	
①		$-u_{z2}$		u_{z4}		u_{z6}	$= 0$
②			$-u_{z3}$		u_{z5}	u_{z5}	$= 0$
③	u_{z1}	$-u_{z2}$	u_{z3}				$= 0$

In der Mascheninzidenzmatrix habe wir $+1$ eingetragen, falls Bezugspfeil von Masche i und Zweigstrom j gleich sind, eine -1 , falls sie gegensinnig verlaufen und eine 0 , falls sich die Pfeile

nicht berühren. In Matrix-Vektorschreibweise lautet dies nun

$$\begin{pmatrix} 0 & -1 & 0 & 1 & 0 & 1 \\ 0 & 0 & -1 & 0 & 1 & 1 \\ 1 & -1 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u_{z1} \\ u_{z2} \\ u_{z3} \\ u_{z4} \\ u_{z5} \\ u_{z6} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

bzw. $Mu_z = 0$ mit der Mascheninzidenzmatrix $M \in \mathbb{R}^{m \times z}$.

Der Zusammenhang zwischen Maschenströmen i_m und Zweigströmen i_z kann ebenfalls in Matrix-Vektorschreibweise formuliert werden, d.h.

$$\begin{pmatrix} i_{z1} \\ i_{z2} \\ i_{z3} \\ i_{z4} \\ i_{z5} \\ i_{z6} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} i_{m1} \\ i_{m2} \\ i_{m3} \end{pmatrix}$$

bzw. $i_z = M^T i_m$.

Der Spannungsverlust entlang eines Zweiges μ ist unter Berücksichtigung einer Zweigspannungsquelle $u_{q\mu}$

$$u_{z\mu} + u_{q\mu} = R_\mu i_\mu.$$

Fast man alle Zweigwiderstände in einer Zweigwiderstandsmatrix R_z zusammen, so lassen sich die Spannungsverluste aller Zweige wiederum in Matrix-Vektorschreibweise formulieren, nämlich

$$\begin{pmatrix} u_{z1} \\ u_{z2} \\ u_{z3} \\ \vdots \end{pmatrix} + \begin{pmatrix} u_{q1} \\ u_{q2} \\ u_{q3} \\ \vdots \end{pmatrix} = \begin{pmatrix} R_{z1} & 0 & 0 & \cdots \\ 0 & R_{z2} & 0 & \\ 0 & 0 & R_{z3} & \\ \vdots & & & \ddots \end{pmatrix} \begin{pmatrix} i_{z1} \\ i_{z2} \\ i_{z3} \\ \vdots \end{pmatrix}$$

bzw. $u_z + u_q = R_z i_z$. Aus der Zweigrelation $R_z i_z = u_z + u_q$ und der Zweigströme $i_z = M^T i_m$ ergibt sich

$$R_z M^T i_m = u_z + u_q.$$

Multipliziert man nun diese Gleichung von links mit M so ergibt sich mit der Maschengleichung $Mu_z = 0$, die Maschenwiderstandsgleichung

$$MR_z M^T i_m = Mu_q.$$

Aus den bestimmbar Größen, der Maschenwiderstandsmatrix $R_m := MR_z M^T$ und der Maschenquellen Spannung $u_m := Mu_q$ kann man nun den Maschenstrom i_m als Lösung des linearen Gleichungssystems

$$R_m i_m = u_m$$

gewinnen. Daraus folgen mit $i_z = M^T i_m$ direkt die Zweigströme und mit $u_z = R_z i_z - u_q$ die Zweigspannungen. In der Praxis wird auch mit induktiven und kapazitiven Widerständen gerechnet, dies bedeutet, dass R_z dann eine Diagonalmatrix mit komplexwertigen Einträgen sein kann. Auch die Anzahl der Zweige und Maschen sind in der Praxis schnell in der Größenordnung $10^2 - 10^4$.

Wir können festhalten, dass in dem Fall einer Diagonalmatrix R_z mit positiven, reellen Diagonaleinträgen, die Matrix $MR_z M^T$ positiv Definit ist. Dass M vollen Rang hat, geht aus den linear unabhängigen Maschen hervor.

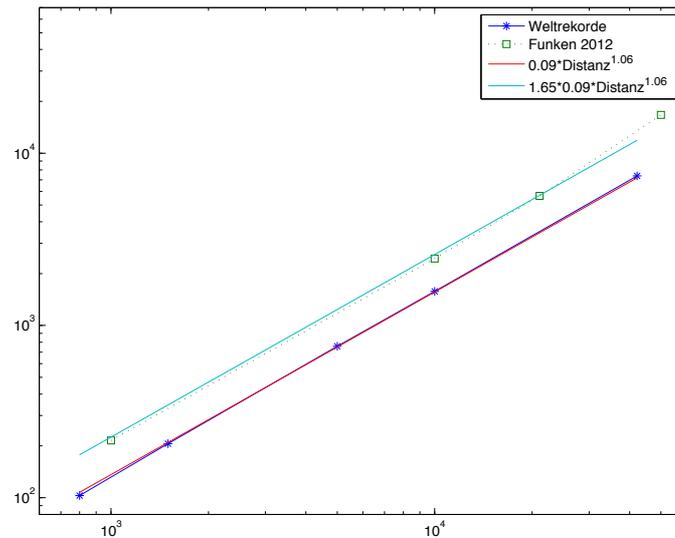


Abb. 1.1: Weltrekorde im Laufen über 800m, 1500m, 5000m, 10km und Marathon und gemessene Werte eines Hobbyläufers.

1.1.2 Ausgleichsproblem

Wie vorhersehbar ist Ihre Bestzeit auf 10.000 m, wenn Sie Ihre 1000m Zeit kennen? Betrachten wir dazu einmal folgende Werte

Distanz	bestehende Rekorde	Hobbyläufer
800	1:43,01 (David Lekuta Rudisha 29.08.2010)	
1000		3:35,00
1500	3:26,00 (Hicham El Guerrouj 14.07.1998)	
5000	12:37,35 (Kenenisa Bekele 31.05.2004)	
10000	26:17,53 (Kenenisa Bekele 26.08.2005)	40:56,00
21098		1:34:11
42195	2:02:57 (Dennis Kimetto 28.09.2014)	3:26:37

und stellen diese in einer Grafik mit doppelt-logarithmischer Skalierung dar, so erhalten wir folgende Grafik.

Scheinbar verhält sich die Funktion der Bestzeit in doppelt-logarithmischer Skalierung annähernd wie eine Gerade, d.h. $Bestzeit(Distanz) = C \cdot Distanz^{1.06}$. Dabei hängt das C von der Person bzw. vom Trainingszustand ab, der Exponent, sprich die Steigung der Geraden in der Grafik, scheint jedoch annähernd gleich zu sein. Wie kann man aber nun die gesuchte Gerade bestmöglich fitten, denn ein exaktes lineares Verhalten wie es wahrscheinlich nicht sein.

Abstrahieren wir die Aufgabenstellung, so kann man folgendes Problem formulieren: Gesucht ist ein Polynom $p(x) = p(x; a_0, \dots, a_n) = a_0 + a_1x + \dots + a_nx^n$, dessen Funktionswerte den kleinsten Quadratsummenabstand zu gegebenen Daten (x_i, f_i) , $i = 1, \dots, m$ haben, das sogenannte Ausgleichspolynom.

Im oben betrachteten Fall suchen wir also eine Ausgleichsgerade. Wie wir in einem späteren Kapitel sehen werden, ergeben sich die Parameter a_0, \dots, a_n gerade als Lösungsvektor a der Normalgleichung

$$W^T W a = W^T f$$

mit

$$W = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & & & & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{pmatrix} \quad \text{und} \quad f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$$

Die Matrix W^T wird auch als Vandermond-Matrix bezeichnet. Einfache Überlegungen zeigen, dass man für eine reguläre Matrix $W^T W$ $n + 1$ paarweise verschiedene Punkte x_i ($i = 1, \dots, m$) benötigt. Gilt dies, so ist $W^T W$ symmetrisch und positiv Definit.

Wo liegen Sie mit Ihren Bestzeiten? Und welche Zeit liefert dies über 10km?

1.2 EINFÜHRUNG, CRAMERSCHE REGEL

Wir beginnen mit dem klassischen Verfahren zur Lösung eines linearen Gleichungssystems (LGS), der Gaußschen³ Eliminationsmethode.

Zu lösen ist ein System von n linearen Gleichungen mit n Unbekannten $x_1, \dots, x_n \in \mathbb{R}$

$$\begin{array}{cccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n & = & b_n \end{array} \quad (1.1)$$

oder kurz

$$Ax = b, \quad (1.2)$$

wobei $A \in \mathbb{R}^{n \times n}$ eine reelle $n \times n$ -Matrix ist und $b, x \in \mathbb{R}^n$ reelle (Spalten-)Vektoren sind.

Wann ist ein lineares Gleichungssystem überhaupt lösbar? Aus der Linearen Algebra (siehe z.B. [Wille]) kennen wir das folgende Resultat, das die Lösbarkeit mit Hilfe der Determinante der Matrix A charakterisiert.

Satz 1.2.1 (Lösbarkeit) Sei $A \in \mathbb{R}^{n \times n}$ mit $\det A \neq 0$ und $b \in \mathbb{R}^n$. Dann existiert genau ein $x \in \mathbb{R}^n$, so dass $Ax = b$.

Falls $\det A \neq 0$, so lässt sich die Lösung $x = A^{-1}b$ mit der **Cramerschen Regel** berechnen, d.h.

$$x_i = \frac{1}{\det A} \begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ a_{21} & \cdots & b_2 & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{nn} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix} = \frac{D_i}{D} \quad (i = 1, \dots, n).$$

Dabei geht die im Zähler stehende Determinante D_i dadurch aus $D = \det A$ hervor, dass man die i -te Spalte der Matrix A durch den Vektor b der rechten Seite ersetzt.

Man beachte hier die Verbindung von Existenz- und Eindeutigkeitsaussage mit dem Rechenverfahren, was einen „guten“ Algorithmus ausmacht. Dieser braucht dabei nicht unbedingt optimal zu sein!

³Carl Friedrich Gauß, 1777 - 1855. Lagrange hatte 1759 die Methode schon vorweggenommen und in China war sie schon vor dem ersten Jahrhundert bekannt. Näheres zu Gauß, Lagrange und weiteren Mathematikern findet man im Internet unter www-groups.dcs.st-andrews.ac.uk/~history.

Die Determinanten von $n \times n$ -Matrizen lassen sich mittels der **Leibnizschen Darstellung** berechnen, d.h.

$$\det A := \sum_{\pi} (-1)^{j(\pi)} a_{1i_1} a_{2i_2} \cdots a_{ni_n},$$

wobei die Summe über alle möglichen $n!$ Permutationen π der Zahlen $1, 2, \dots, n$ zu berechnen ist. Der Wert des Ausdrucks $(-1)^{j(\pi)}$ ergibt sich aus der Anzahl $j(\pi)$ der Inversionen der Permutation $\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$.

Bemerkung 1.2.2 (Rechenoperationen) *Im Folgenden werden die Verknüpfungen Multiplikation, Addition, Division und Subtraktion in ihrem Rechenaufwand nicht unterschieden und unter dem Begriff **Gleitkommaoperation** zusammengefasst (1 Gleitkommaoperation $\simeq 1$ FLOP⁴). Systematische Multiplikationen mit ± 1 bleiben im Allgemeinen unberücksichtigt. Anderweitige Operationen wie z.B. Wurzelziehen werden gesondert betrachtet.*

Satz 1.2.3 (Aufwand Leibnizsche Darstellung) *Sei $A \in \mathbb{R}^{n \times n}$. Der Aufwand zur Berechnung von $\det A$ mit der Leibnizschen Darstellung, d.h. als Summe über alle Permutationen der Menge $\{1, \dots, n\}$, beträgt*

$$\text{FLOP}(\det A) = n n! - 1.$$

Beweis. Der Aufwand zur Berechnung von $\det A$ für $A \in \mathbb{R}^{n \times n}$ in der Leibnizschen Darstellung (unter Vernachlässigung der Multiplikation mit $(-1)^{j(\pi)}$) ergibt sich wie folgt:

$$\begin{aligned} \text{FLOP}(\det A) &= \text{'' } (n! - 1) \text{ Additionen } + n! \text{ Produkte mit } n \text{ Faktoren ''} \\ &= n! - 1 + n!(n - 1) = n n! - 1. \end{aligned}$$

Damit ist die Aussage des Satzes bewiesen. □

Alternativ lässt sich die Determinante rekursiv mit Hilfe des **Laplaceschen Entwicklungssatzes** berechnen. Dieser lautet für eine quadratische Matrix $A \in \mathbb{R}^{n \times n}$

$$\det A = \sum_{j=1}^n (-1)^{1+j} a_{1j} \det(A_{1j}),$$

wobei A_{1j} diejenige Matrix ist, die aus A durch Streichen der ersten Zeile und der j -ten Spalte entsteht.

Satz 1.2.4 (Aufwand Laplacescher Entwicklungssatz) *Sei $A \in \mathbb{R}^{n \times n}$ ($n \geq 2$). Der Aufwand zur Berechnung von $\det A$ mit dem Laplaceschen Entwicklungssatz beträgt*

$$\text{FLOP}(\det A) = \sum_{k=0}^n \frac{n!}{k!} - 2 < e n! - 2.$$

Beweis. Die Ungleichung ist klar. Die Gleichung lässt sich induktiv beweisen:

Induktionsanfang ($n = 2$): Die Determinante der Matrix $A = (a_{ij}) \in \mathbb{R}^{2 \times 2}$ lautet

$$\det(A) = a_{11}a_{22} - a_{21}a_{12},$$

⁴FLOP ist nicht zu verwechseln mit FLOP/s oder flops (floating point operations per second), welches als Maßeinheit für die Geschwindigkeit von Computersystemen verwendet wird.

d.h. der Aufwand beträgt 2 Multiplikationen und 1 Addition, also

$$\text{FLOP}(\det(\mathbb{R}^{2 \times 2})) = 3 = \frac{2!}{0!} + \frac{2!}{1!} + \frac{2!}{2!} - 2.$$

Induktionsschritt ($n \rightarrow n + 1$): Für $n \geq 2$ gilt

$$\begin{aligned} \text{FLOP}(\det(\mathbb{R}^{(n+1) \times (n+1)})) &= \text{"}n \text{ Additionen} + (n+1) \text{ Multiplikationen} \\ &\quad + (n+1) \text{ Berechnungen von } \det(\mathbb{R}^{n \times n}\text{"} \\ &= n + (n+1) + (n+1) \cdot \text{FLOP}(\det(\mathbb{R}^{n \times n})) \\ &\stackrel{\text{IH}}{=} 2n + 1 + (n+1) \cdot \left(\sum_{k=0}^n \frac{n!}{k!} - 2 \right) \\ &= 2n + 1 + \sum_{k=0}^n \frac{(n+1)!}{k!} - 2(n+1) \\ &= \sum_{k=0}^n \frac{(n+1)!}{k!} - 1 \\ &= \sum_{k=0}^{n+1} \frac{(n+1)!}{k!} - 2. \end{aligned}$$

□

Beispiel 1.2.5 Ein einfaches Beispiel soll zeigen, dass man die Determinante überhaupt nur für sehr kleine allgemeine Matrizen der Dimension $n \ll 23$ mit dem Laplaceschen Entwicklungssatz berechnen kann.

Ein Jahr hat $365 \cdot 24 \cdot 60 \cdot 60 \approx 3 \cdot 10^7$ Sekunden. Geht man nun von einem schnellen Rechner⁵ mit 180 TFlops aus, so benötigt man zur Berechnung der Determinante einer 20×20 -Matrix mit dem Laplaceschen Entwicklungssatz

$$\frac{\text{Anzahl der Operationen}}{\text{Gleitkommaoperationen pro Sekunde}} [s] = \frac{\sum_{k=0}^{20} \frac{20!}{k!} - 2}{180 \cdot 10^{12}} [s] \approx 36741 [s] \approx 10.2 [h]$$

bzw. für eine 23×23 -Matrix

$$\frac{\sum_{k=0}^{23} \frac{23!}{k!} - 2}{180 \cdot 10^{12}} [s] \approx 3.904 \cdot 10^8 [s] \approx 12.38 \text{ Jahre}.$$

Bemerkung 1.2.6 Bei der Cramerschen Regel ist zum Lösen eines linearen Gleichungssystems $Ax = b$ mit $A \in \mathbb{R}^{n \times n}$ und $b \in \mathbb{R}^n$ die Berechnung von $n + 1$ Determinanten und n Quotienten notwendig. Der Aufwand zur Lösung eines linearen Gleichungssystems lässt sich somit zusammenfassen, wobei wir auf die Komplexität des Gauß-Verfahrens (siehe Seite 14) erst später eingehen werden.

Cramersche Regel		Gauß-Elimination
Leibniz	Laplace	
$n(n+1)! - 1$	$\sum_{k=0}^n \frac{(n+1)!}{k!} - n - 2$	$\frac{4n^3 + 9n^2 - n}{6}$

⁵Der weltweit auf Platz sechs rangierende und zugleich schnellste europäische Supercomputer JUGENE am nordrhein-westfälischen Forschungszentrum Jülich hat 180 TFlops (Stand Okt. 2008). Ein aktueller PC hat eine Leistung von 1-10 GFlops. T = Tera = 10^{12} , G = Giga = 10^9 .



Bemerkung 1.2.7 Für das Lösen eines LGS mit 20 Unbekannten mit der Cramerschen Regel und dem Laplaceschen Entwicklungssatz benötigt man auf einem der schnellsten Rechner mehr als eine Woche, d.h. $21 \cdot 10.2 [h] = 214.2 [h] = 8.925 [d]$ (20 verschiedene Determinanten im Zähler und eine im Nenner, vgl. Beispiel 1.2.5). Ein System mit 23 Unbekannten ist mit einem heutigen Supercomputer und der Cramerschen Regel nicht in einem Menschenleben zu lösen.

Beispiel 1.2.8 (Vergleich Rechenaufwand) Aufwand zum Lösen eines linearen Gleichungssystems $Ax = b$ via Cramerscher Regel und Gauß-Verfahren.

Für einige n sei die Anzahl der notwendigen FLOPs wiedergegeben.

n	Cramer/Leibniz $= n(n+1)! - 1$	Cramer/Laplace $= \sum_{k=0}^n \frac{(n+1)!}{k!} - n - 2$	Gauß $= (4n^3 + 9n^2 - n)/6$
$n = 2$	11	11	11
$n = 3$	71	59	31
$n = 4$	479	319	66
$n = 5$	3599	1949	120
$n = 8$	2903039	986399	436
$n = 10$	399167999	108505099	815

Im Folgenden werden wir nun Verfahren beschreiben, die bereits für $n \geq 3$ effektiver als die Cramersche Regel sind. Mit Hilfe dieser Verfahren lassen sich auch Determinanten in polynomieller Zeit bestimmen. Daher wird die Cramersche Regel im Allgemeinen nur für $n = 2, 3$ verwendet. Der Vorteil der Cramerschen Regel ist jedoch die explizite Schreibweise, d.h. sie ist eine Formel für alle Fälle. Man spart sich bei ähnlichem Rechenaufwand (d.h. $n = 2, 3$) aufwendige Fallunterscheidungen (z.B. Pivotwahl) im Programm.

1.3 GESTAFFELTE SYSTEME

Betrachten wir zuerst besonders einfach zu lösende Spezialfälle. Am einfachsten ist sicherlich der Fall einer diagonalen Matrix A .

Diagonalmatrix

Man bezeichnet eine Matrix $A = (a_{ij})$ als Diagonalmatrix, falls $a_{ij} = 0$ für $i \neq j$ gilt. Häufig schreibt man eine Diagonalmatrix A mit Diagonaleinträgen a_{11}, \dots, a_{nn} auch einfach $A = \text{diag}(a_{11}, \dots, a_{nn})$. Die Anwendung der Inversen einer Diagonalmatrix A mit Diagonalelementen $a_{ii} \neq 0$ auf einen Vektor b lässt sich als Pseudocode wie folgt schreiben: